

Protection for laboratory animals?

The public regulation of the use made of laboratory animals is necessarily an awkward compromise between the interests of research and medicine on the one hand and the interests of the wider population on the other. The basis of all animal regulations is also philosophically fuzzy. Questions such as whether animals "feel pain" in the sense in which the term applies to those animals (human beings only?) equipped with consciousness are undecided. And while it is widely accepted, even in hard-line laboratories, that some restraint is needed, neither the scientific community nor those who advocate increasing restraint on the uses of laboratory animals are clear about the basis on which regulations should be drawn. The result is widespread incoherence, in which the extremists among the animal-lovers have taken to downright irrationality — in Britain, burglary and the theft of laboratory animals and the daubing of paint on the houses of well-known scientists often mistaken for vivisectionists (including, unaccountably, Sir Richard Doll). The danger now is that the new regulations for the protection of laboratory animals in prospect in several countries will be drawn up with as little concern for logic as were the regulations they will replace.

Events in Britain, renowned for its differential love for certain species (mostly those to which cute diminutives such as "pussy" or "bunny" can be applied) are potentially a guide to what may happen elsewhere. In spite of a vigorous campaign against "vivisection" for the past century, the laboratory uses of animals are still regulated by the Cruelty to Animals Act of 1876. The act is administered by the Home Office, which issues licences to experimenters and maintains a staff of fifteen inspectors to monitor observance of the conditions attached to licences. In the past few years, however, even the government has come to accept that the time has come for reform. Since 1979, the issue has been alive in the British Parliament. In May, the Council of Europe finished drafting a convention that is yet another spur to change. Meanwhile, the Home Office has asked its Advisory Committee on Animal Experiments to suggest what should be done — and has thanked it by saying that "its report does not provide a definitive basis for . . . legislation".

The need for change is obvious enough, and put clearly in the advisory committee's report (which may be had free of charge from the Home Office). Since 1876, the uses of laboratory animals have been widened enormously; experimental psychology was not then foreseen, for example. Moreover, the act does not apply to the use of animals other than for "experiments" nor to the establishments in which laboratory animals are bred or kept. The advisory committee (whose chairman is Dr Mary Warnock) is rightly scornful of the requirement that an application for a licence should be countersigned by the president of one of nine nominated learned societies; public administration should be rid of the procedure of the rubber stamp. It rightly says that the most important change since 1876 has been the scale on which laboratory animals are now used. But, while predictably a model of clarity, it falls short of acknowledging the consequences for seemingly and effective regulation of the growth of the numbers of animals in laboratories in the past century.

But how is this to be accomplished? The first need is that there should be a fuller understanding of the diversity of the case against the use of laboratory animals. Existing legislation, in Britain and elsewhere, is based on the assumption that procedures that would cause pain (were it not for anaesthesia) are justifiable only by their objectives as assessed by human beings — better medicine (or even better veterinary practice), better under-

standing or better teaching. Thus the 1876 Act requires that experiments involving pain must be carried out under anaesthesia (and the Home Secretary has not waived this regulation for the past half-century). The moderate case against the existing regulations involves the range of their application — only vertebrates? even embryonic forms such as tadpoles? — but also asks that experiments and other procedures involving animals should be kept to a minimum. Much of this case is conceded by those working with laboratory animals. The abiding irony is that the more stringent testing of medicines, and attempts to spot occupational carcinogens in advance, entail a potentially very large increase of the scale on which experiments with animals are conducted. Yet disputes persist, and in principle cannot be resolved objectively. People differ in their necessarily subjective assessment of what constitutes pain, or excessive pain, in a member of some other species. Even the version of the Laboratory Animals Protection Bill first introduced in the House of Lords by Lord Halsbury, which has since been amended (with the government's approval) by a select committee, would not avoid such difficulties.

More extreme arguments against laboratory experiments with animals are not, however, simply to be scorned. Quite apart from disputes of a quantitative kind about the degree of human benefit expected from particular uses of laboratory animals, there are many who hold that other than human animals deserve equal consideration and respect, with the result that no laboratory procedures with animals are justifiable. This view is not exclusively religious, even if a central tenet of Buddhism. Many of those who hold it have been driven to their radical conclusion by ecological arguments (or even misunderstandings), or by a generalized conviction, misguided though it may be, that the benefits of all scientific work are overrated.

Those who dissent from these opinions, presumably including all those who work with laboratory animals, must nevertheless not fall into the trap of supposing that they have no weight. They will certainly influence the British Home Office in drafting legislation — it may already be significant that the British delegation to the Council of Europe is kicking up a fuss about the provision (Article 8) of the draft convention that would permit licensed experiments causing pain without anaesthesia even though the convention is only permissive. The trouble is that there is hardly ever any meeting between the radically opposed views of those who would continue to use laboratory animals and those who would stop them altogether.

But why should the objectives of medicine and research be compromised by forces of unreason? That will be the complaint. Up to a point, it is also the basis of the 1876 Act, which assumed that it was for the government to strike a wise balance between opposing interests and then to administer the regulations for the rest of time. This process is too static, too old-fashioned and provides too few opportunities for the arguments about the proper use of laboratory animals to be engaged and, with luck, resolved. Yet even the intelligent report of the advisory committee falls into the trap of supposing that the present system needs to be modified only to the extent that there should be a more powerful committee, able to advise on the interpretation of the law in novel circumstances.

More daring is called for. The polarization between the paint-daubers and the scientific community will not be resolved by a new law in the old mould, however well considered. What is needed is a mechanism that will repeatedly bring opposing views



on the use of animals together in circumstances in which there are practical and not merely grand philosophical questions to be decided. So why not build into a new system of regulation a device that has worked well in the regulation of genetic manipulation and of experimental procedures with human subjects — a committee at every important centre at which animals are used that would be empowered to sanction (or not to sanction) proposed uses of laboratory animals? Such a mechanism would be more than an exercise in public relations, for it could help to define if not to answer the philosophical questions that at present ensure that all legislations seems arbitrary to those most concerned. It need not impede the conduct of research nor rob the government of the political responsibility that it seems to cherish. And, if well run, such a system, while requiring of some scientists an extra committee meeting, could simplify and rationalize the present system of licensing.

Away with Kondratiev

Little has been heard of the Centre for Technical Change since it was first set up by the Leverhulme Trust and a consortium of British research councils just over a year ago. The delay, a little at odds with the declared objectives of understanding what is wrong with the British economy, is forgivable in that, as always, the recruitment of staff for a purpose-built organization is necessarily slow. Some hint of which way the wind is blowing can, however, be gleaned from the paper delivered earlier this week to the Institute of Ceramics by the director of the centre, Sir Bruce Williams. His theme was the question of how to strike a balance between technical innovation, economic growth and leisure either enforced — as unemployment — or arranged as by means of education (which delays entry to the labour force), shorter working weeks or earlier retirement. His paper should be read in conjunction with a study, 'Innovation and Efficiency' by Haustein, Maier and Uhlmann, now published in the institute's series of *IIASA Reports* (3, 309; 1981).

Both arguments start by distinguishing between product and process innovations — respectively the innovations that introduce new products to the market and those which result either in the improvement of existing products or more efficient ways of making them. Of necessity, product innovations increase jobs, and process innovations decrease them. Williams refers to the Kondratiev cycle, the forty-to-fifty year cycle in economic activity attributed by Schumpeter to the coincidence of innovations of all kinds at certain periods of time. Haustein *et al.* have carried out a study of the decline of productivity in industrialized countries over the past thirty years (in both "planned" and capitalist economies, Japan included) and, concluding that there is indeed a cycle in the "efficiency" of the process of innovation, unremarkably suggest that the success of an innovation is determined by such considerations as market potential, the organization of the firm or organization concerned and "know-how", whatever that may be.

Williams, as might be expected, is more pragmatic. This may be a time of downturn in the Kondratiev cycle, and thus a time when the opportunity for jobs is also declining. But, with luck, there will eventually be an upturn. In the meantime, it might help if governments were to concentrate support for research and development on job-creating innovations. All this is sensible enough. But can either analysis be correct? Even the past few years have not been short on product innovations. The benefits of microelectronics have been widely touted, yet the benefits have not been quickly spread about because potential customers have not been able to afford them. Part of the trouble is that the manufacture of microelectronic equipment is labour intensive, which means that potential purchasers of, say, a home computer must also undertake to provide somebody else's food and drink. The same, so to speak, is even more true in telecommunications, where the cost of installation often far outweighs the cost of manufacture. It is also manifest that the growth of most economies in the past thirty years has been sustained by the

improvement of agricultural productivity. So perhaps the way to avoid the Kondratiev cycle is to invest in process and not product innovation, contrary to expectation. Enabling the market to function is at least half the battle.

Back to the Pleistocene

Western Europe is plainly in for another bout of preoccupation with the building of a tunnel beneath the English Channel or La Manche, according to where you live. Observers of last week's meeting between the British Prime Minister, Mrs Margaret Thatcher, and the President of France, M. Francois Mitterrand, seem unaccountably to have been at a loss to understand their cordiality (yet both are well-mannered people) and their decision to resurrect once again this ancient technological project, so long in the tooth that it was first abandoned when the British took fright that Napoleon might march an invading army beneath the sea. Last week's agreement to carry out yet another feasibility study of the tunnel is, however, not at all surprising. The two politicians may differ in their politics but they share a common preoccupation with high and rising unemployment and the knowledge that its amelioration by the spending of domestic resources would entail that the printing presses would have to work harder than at present and that the consequences would be inflationary. Among megaprojects, however, the Channel Tunnel has the advantage that it might be possible to persuade the rest of the European Community to pay for a substantial part of the cost, on the grounds that it would assist the flow of trade within Europe as a whole. Yet whatever the European Community contributes, the chances are high that the immediate benefits would accrue in Britain and France. So much of the feasibility study will be among the moneybags of Brussels.

Even if the immediate objective is economic, however, it will be in the public interest, both in Britain and in France, that technical considerations should not be overlooked. Some attention should also be paid to the reasons why previous proposals have collapsed. In 1975, for example, the proposed rail tunnel between Britain and France foundered on the gloomy calculations by British Rail of the cost of providing a railhead in Kent and a terminus in central London. The correctness of the decision was reinforced by the widespread disbelief that the railways would be able to load special trains with vehicles at a rate of one every few minutes and by the no doubt cynical conviction that to give the British and French railways a monopoly of wheeled communication between the two countries would merely provide them with a permanent means of recovering the cost of uneconomic services elsewhere. The lesson is that if there is to be a direct link between Britain and France, it should not be exclusively a rail link.

With the passage of time since 1975, other candidate proposals for a direct link have emerged. British Steel and other European steel-making concerns are understandably keen to build a bridge most of the way across, using sections of immersed-tube tunnel in the shipping lanes. Unfortunately, none of these plans takes full account of the suitability of the Channel (or of La Manche) as the site for a feasible megaproject — the fact that it is nowhere deeper than fifty metres, with an average depth of some thirty-five metres. In such circumstances, dam-building is by no means out of court. Indeed, with a little ingenuity, the civil engineers could arrange to provide not merely a means of walking from Britain to France but also of providing the whole of Europe with electricity. To the complaint that a dam would interfere with shipping there is the quite proper riposte that most of the ships are there only because the Channel (or La Manche) has taken the place of the late Pleistocene land-bridge. Industrial consortia wishing to pursue this option should consult the report of the competition to dam the Channel organized by *Nature* in 1968 (220, 1168). They should reflect that no harm will come from thinking big, for the new feasibility study, like its predecessors, is likely to conclude that building a link between Britain and France is no longer economically feasible even though it would have paid off handsomely if the decision had been taken at some earlier stage.

Time runs out for space telescope

Guidance snags licked, others still hidden

Washington

Many fingers were crossed at a groundbreaking ceremony held at Johns Hopkins University in Baltimore, Maryland, on Monday to mark the beginning of construction work on the Space Telescope Science Institute. The construction of the Large Space Telescope itself, due to be launched from the space shuttle early in 1985, seems just about on schedule and no major technical hurdles are foreseen.

Those who attended a programme review at the National Aeronautics and Space Administration (NASA)'s Marshall Space Flight Center two weeks ago, however, came away nervous that unanticipated problems could disrupt the delicate equilibrium in which the whole project seems to be balanced.

The Space Telescope Science Institute will be the principal observation and data analysis centre for the telescope, which will be placed in orbit 500 kilometres above the Earth's surface. It is being built on Johns Hopkins' Homewood Campus by the Association of Universities for Research in Astronomy under an initial five-year, \$24 million contract with NASA. Dr Riccardo Giacconi has been named director of the institute, which will eventually have a staff of about 150, including between 30 and 40 astronomers from the United States and abroad.

The good news on the telescope is that the two mirrors which make up its Ritchey-Chrétien optical system are almost completed and are said to work beyond expectations. But there is less certainty about the eventual outcome of the instrumentation being developed to maintain the direction and stability of the telescope in orbit.

Initial designs for the fine-guidance sensor called for an image dissector able to capture a selected guide star combined with a fixed interferometer to compare the separate images of the star produced. When this proved to be a cumbersome arrangement Perkin-Elmer, the principal contractor for the Optical Telescope Assembly, put forward a more streamlined design which incorporated the moving parts into the interferometer. Once the star selector had found the guide star, the whole fine-guidance system would be locked in place, and directioning control transferred to the gyroscopes which are used to position the space support module within which the optical telescope assembly is contained. The scheme was calculated to operate with a high degree of accuracy.

Unfortunately it was realized at a late stage that, as the fine guidance sensor had a limited near range, by locking it in place the system would be unable to compensate for any sudden torque produced in the system as a whole in reaction to the movement of some other component if the result was a linear displacement of more than about 0.015 arc seconds.

Following this "pointing crisis", the whole system has been redesigned. New feedback has been introduced by keeping the star selector in the interferometer working in an active mode to keep the telescope fixed in its predetermined direction. Several parts of the five scientific



Primary mirror — looking good

instruments being carried are also being altered to reduce the effects of their movement on the rest of the vehicle.

However, the additional movement in the pointing system could still cause problems if either the fine-guidance sensor

or the gyroscopes create some unanticipated resonance in the support structure, something that will only be discovered during the pre-launch phase when a modal survey will be carried out.

The alterations will also mean a slight degradation in the telescope's expected performance. This will still be well within the minimum performance included in the original design specifications; but it will not, at least initially, be as high as it could have been if more time (and development funds) had been available.

Under the original Perkin-Elmer design, for example, the telescope was expected to be able to maintain an accuracy of 0.007 arc seconds over extended periods of time. The new system is designed to an accuracy of 0.012 arc seconds although it is hoped that in practice this can be raised to 0.009 arc seconds. Although much higher than most astronomers need, this will, for example, result in a loss of some image quality in the ultra violet range with the faint-image camera being built by the European Space Agency.

According to project scientist Dr Robert O'Dell, however, the main problems faced by the space telescope are not technical but financial. The development programme — costing about \$120 million a year — is at the stage when delays can have the most significant impact on overall costs.

Late last year NASA reviewed the programme in view of problems that seemed to be emerging at Perkin-Elmer. This set the launch date back by eighteen months to early 1985, and the total predicted cost was increased by about \$150 million in 1982 dollars to between \$700 and \$750 million.

From the frying-pan into the fire?

The British Prime Minister, Mrs Margaret Thatcher, announced earlier this week that the Secretary of State for Education and Science since May 1979, Mr Mark Carlisle, has been replaced by Sir Keith Joseph, Secretary of State for Industry during the same period. The changes are part of a substantial reorganization of the British Cabinet.

Mr Carlisle will now become an ordinary Member of Parliament. Sir Keith will remain a member of the Cabinet, in which he has been one of the principal advocates of monetarist policies since the election of the present government, but the post of Secretary of State for Education and Science is usually reckoned to carry with it less political clout than that at industry.

Even so, and in spite of Sir Keith's reputation for agonizing hard and long (and sometimes too long) before making a decision, the arrival of a new minister at the department will be welcomed in higher education. Mr Carlisle's uncompleted business includes the need to respond to the demand of the House of Commons Select Committee on Education for a sight of the

letters written to him by the chairman of the University Grants committee, Dr Edward Parkes, and the need to respond to the demand by the Committee of Vice-Chancellors and Principals that funds should be made available to pay for the cost of redundancies among academics forced on the universities in July. The arrival at the Department of Education and Science of a secretary of state free to disown his predecessor's decisions will seem to many universities to be an opportunity for reprieve.

The reshuffle may, however, have come too late for some universities. Those which have already drawn up plans for reducing academic payrolls by compulsory redundancy may find that their compliance is taken as a sign of weakness. The rest are likely now argue that a slower contraction of the university system than decreed will in the long run be more economical.

The Space Telescope remains at the top of NASA's space science priority list. Proposed to Congress by President Carter in 1977, it is the last "new start" to have escaped budget cuts, and its projected launch date has remained fixed on the

space shuttle schedule in recent months while others have gradually slipped back.

So far, everyone is being optimistic that the 1985 or possibly an earlier launch date will be met. This will be particularly important to planetary scientists who will need observations from the space telescope early in 1985 to prepare the imaging sequence for the Voyager 2 spacecraft as it approaches its encounter with Uranus on 24 January 1986.

If the programme does slip, NASA could be faced with a rapidly escalating bill as technical crews and scientific facilities are kept standing idle. The major nightmare is that as with the shuttle, budget increases caused by last-minute delays could have a considerable impact on other parts of the space science programme. **David Dickson**

DNA guidelines

Voluntary code

Washington

In a move that is expected to precipitate the withdrawal of government controls on recombinant DNA research, the National Institutes of Health (NIH) are being provisionally recommended by their Recombinant DNA Advisory Committee (RAC) to turn the safety guidelines which must at present be followed by all those receiving federal funds for such research into a voluntary code of practice.

A resolution passed at a public meeting of the committee in Bethesda, Maryland, last Thursday recommended that, rather than expecting research workers to obey a special set of rules for experiments using recombinant DNA, in general the safety measures to be taken should be comparable with those suggested by, for example, the Center for Disease Control (CDC) for research involving the microorganisms in question.

Changes suggested by RAC in the existing guidelines, first introduced by NIH in 1976, would keep the committee in existence as a forum for public discussion of potential safety problems. Several committee members acknowledged that its function would be primarily social and political, designed to deflect any future local opposition to recombinant DNA research.

Individual universities and research institutions, however, would be free to decide how they administered the guidelines, and would no longer be required by NIH to do this through a formally-established Institutional Biosafety Committee (IBC) with a prescribed membership, as at present. Nor would most experiments require prior approval by IBCs or their equivalent.

RAC's actions followed suggestions by Dr Allan Campbell of Stanford University and Dr David Baltimore of the Massachusetts Institute of Technology, for a substantial easing up both in the containment criteria to be used for different types of experiments, and in the administrative processes required to see that these

were followed (*Nature* 291, 3; 1981). These recommendations were passed to a subcommittee which presented its report to the full committee last week. In general, the members of the subcommittee agreed that there was a good scientific case for a significant reduction in the containment levels for all experiments apart from those known to present a particular type of danger.

In its agreed statement, which will be published in the *Federal Register* for public comment before being given final consideration by RAC, the full committee suggests that where physical containment levels applicable to non-recombinant DNA experiments exist for either the host or the vector, for example those under revision by CDC, "recombinant DNA experiments should be carried out at containment levels at least as high as those recommended for non-recombinant DNA experiments".

Extensive discussion took place about which types of experiment should be prohibited (or under a voluntary code of practice, should not be carried out) without explicit consideration. Committee members voted that it should no longer be necessary to suggest this either for large-scale experiments using more than 10 litres of culture, or for experiments which involved the release of genetically-altered microorganisms into the environment.

However, the committee also agreed that special consideration should still be given to experiments which involve the transfer of drug-resistance traits, as well as work with known toxins, for which specific containment levels would still be recommended.

There was less agreement on how much of an administrative apparatus should be kept in place. Professor Norton Zinder of the Rockefeller University made a strong plea to the committee that, because the scientific basis for imposing strict regulations on research workers seemed to be evaporating, the RAC should consider phasing itself out of existence.

Dr Baltimore responded by saying he felt RAC could continue to play an important role. In the past, for example, congressional legislators had been able to use the existence of the committee to ward off demands for regulatory legislation. In the same way he said, in communities such as Boston where the activities of emerging genetic engineering companies has once again stimulated a fierce public debate over the potential health and environmental risks, the absence of a national focus could stimulate demands for increased local surveillance and control.

The committee as a whole seemed eager to get on with dismantling the regulatory structure that has been built up by NIH. By a vote of 16 members to three, with one abstention, it agreed to a statement recommending revisions to the guidelines that would eliminate their mandatory element, remove the now-redundant section suggesting voluntary compliance by private industry, remove several

categories from the list of prohibitions (which would be referred to as "admonitions"), and merely include the statement that "adherence to these standards by all laboratories using recombinant DNA is recommended". **David Dickson**

Plant genetics

Head-hunt threat

Britain could lose the fruits of research of genetic manipulation with plant material unless steps are taken soon to exploit techniques now being developed. That is the concern of the British Technology Group (formed last month from the former National Research Development Corporation and the National Enterprise Board). The group is especially concerned about commercially promising developments within the Agricultural Research Council's special programme.

The chief concern is that the council's programme has been so successful as to arouse the interest of biotechnology companies in the United States. Representatives from several companies, including Advanced Genetic Science from Greenwich, Connecticut, and Boston-based Genetics International, have been head-hunting in council institutes. The British Technology Group fears that the US companies will tempt away people and ideas before British industry can make an alternative investment.

Most efforts so far have comprised job offers in the United States and money for fellowships in Britain, but some companies are said also to be considering setting up laboratories in Britain. Joint projects with the Agricultural Research Council are ruled out because of the understanding that the British Technology Group should have first refusal on commercially promising developments by research council scientists.

All parties agree that there is no immediate threat of an exodus of British plant geneticists, but each recognizes that a British initiative is needed fairly urgently if the lure is not to become too great. One of the possibilities being looked at is for a company similar to Celltech; one fear is that even such a company might rob the Agricultural Research Council of staff.

Like Celltech, the new company might begin by supplying monoclonal antibodies for veterinary use, among other things. Research under the Agricultural Research Council's genetic manipulation programme is too long-term, however, to be of immediate application. It includes expressing nitrogen-fixing genes in chloroplast cells, expressing chloroplast genes in *Escherichia coli* and the cloning of genes specific to barley storage protein. Recent progress in these and other areas, however, seems to have persuaded the American companies that they are worth investing in even now.

Judy Redfearn

Polish education

Solidarity solid

Last week's Congress of Solidarity in Gdansk called on the Polish government to introduce a bill on higher education that would be acceptable to members of the academic and learned professions. This reflects a growing concern in Polish academic circles that the very considerable gains in academic autonomy won during the past year might be reduced when the negotiations between government and academics are codified in a new law.

The gains range from the right of the senate of a university to elect its own rector and dean through student participation in university government to the abolition of censorship of imported publications needed for academic study. It has also been agreed that compulsory Russian for undergraduates should be replaced by the choice of a foreign language.

So far, however, these concessions have been made piecemeal. The most comprehensive attempt to codify them — the Lodz Accords, which ended a month of student unrest last February — was expressed in terms so vague that there were considerable doubts in academic circles as to how binding they might be.

During the past six weeks, there have been several hints that the Polish authorities, while not actually reneging on their promises, are trying to draw back from the most radical demands. From the new censorship law, passed at the end of July, there seems to have been struck out the promised clause that a returning traveller could import for his or her own use one copy of any foreign publication, however controversial.

The new Independent Students' Association, legitimized by the Lodz Accords, is also concerned about apparent attempts by the authorities to whittle down student power. These include the deportation, in February, of a Czech student, Lenka Cvrckova, who had played a major part in organizing the Lodz protests. (She is now in gaol in Prague awaiting retrial on charges of passport irregularities after an acquittal by a lower court was overruled.)

The Ministry of Science, Higher Education and Technology has also refused to permit legal recognition to the student unions of Poland's three main ethnic minorities — Lithuanians, Byelorussians and Ukrainians — and it is suggested that the promised 33 per cent student representation on the senates of universities may be reduced to 20 per cent. Spokespersons for the students' association stress that during the coming academic year there are unlikely to be major clashes between students and the university governing bodies — since the rectors and deans have already been elected by the new procedures. The possibility of clashes between students and the ministry

is, however, far from remote.

So sensitive is the situation in some universities, apparently, that there have been suggestions that the opening of the academic year may be postponed — ostensibly on account of the economic crisis. Of greater significance, however, is the meeting last week between the new Minister of Science, Higher Education and Technology, Dr Jerzy Nawrocki, and representatives of the Union of Polish Teachers. This body, which includes university lecturers as well as schoolteachers, is not allied with Solidarity but is one of the old pre-1980 party-linked unions. Like Solidarity, however, this union considers that the government's draft bill on higher education shows a number of "negative deviations" from the version prepared by the "social codification commission". Minister Nawrocki has replied that apart from a few editorial changes, all the changes were "indispensable" modifications, reflecting the existing legal order and the powers of the state organs. The Union of Polish Teachers is unimpressed, and has called for further talks.

Vera Rich

UK atomic energy

Research reaction

The United Kingdom Atomic Energy Authority last year stepped up research on pressurized water and advanced gas-cooled reactors, partly in reaction to the government's commitment to a substantial programme of building thermal nuclear power stations. But fast reactor research consumed the lion's share — £82 million — of the £114 million 1980–81 nuclear research and development budget.

The authority is, however, cheerful about progress in the fast reactor division. Dr Walter Marshall, the authority's new chairman, announced last week that a new conceptual design for a commercial demonstration fast reactor will be ready shortly. The design, the latest of several drawn up by the National Nuclear Corporation in conjunction with the authority, is said to be "compact and elegant", and also cheaper than previous designs. Indeed, the authority estimates that a demonstration plant will cost only about 20 per cent more than a conventional advanced gas-cooled reactor, probably making it cheaper than France's Superphénix.

Dr Marshall seems, nevertheless, less concerned than his predecessor, Sir John Hill, that the government should agree to finance such a reactor now. The implication is that the new commitment to a thermal reactor programme will keep the authority and the rest of the nuclear industry busy for some time. Dr Marshall would prefer to wait until after the public inquiry into the Sizewell pressurized water reactor, but would appreciate a

statement of intent much sooner.

The fear that a lengthy delay in making such a commitment could jeopardize negotiations over international collaboration, in particular with the United States and France, seems to have receded. Dr Marshall, who has been talking to the US and French administrations, is optimistic that some arrangement will be possible about collaboration on fast reactors.

The authority has increased research into the safety of pressurized water reactors, in particular into the integrity of pressure vessels, the hazards of loss of coolant accidents and the impact of a major, but improbable, accident which, it says, would not release as much radioactivity as had previously been thought. The authority has also proposed that an Inspection Validation Centre be set up to train inspectors of pressure vessels, but has yet to make a formal application for the facility to the government.

Dr Marshall, who was speaking during the presentation of the authority's 1980–81 annual report, did not, however, have anything to say about the progress of the task force established earlier this year, under his chairmanship, to rethink the National Nuclear Corporation's design of a British pressurized water reactor. That design had become overcomplicated in an attempt to incorporate extra safety features.

The authority's expenditure during 1980–81 was slightly more in real terms than in the previous year. Total expenditure was about £331 million, of which £187 million came from parliamentary grant, the rest from income earned through services to industry.

Judy Redfearn

Computer memories

Bubbles burst

Washington

At the beginning of the year bubble computer memories, first developed by Bell Laboratories in the 1960s and introduced to the market four years ago, seemed poised for take-off. Although sales up to that point had been slow, they were predicted to rise sharply to about \$40 million in 1981, and four large US companies — Texas Instruments, National Semiconductors Corporation, Rockwell International and Intel — were publicizing plans to compete for a substantial share.

Now only Intel remains. National Semiconductors has announced its intention to withdraw from the bubble memory business, a decision that followed similar announcements by Texas Instruments and Rockwell earlier in the year. Intel remains confident that it can still find specialist applications in which bubble memories can meet its initial expectations. But grandiose schemes for their becoming a major competitor to conventional disk storage and silicon chips now seem unlikely to materialize.

National Semiconductors' decision was

not based solely on its experience with bubble technology. The company, which is the second largest producer of integrated circuits in the United States, is going through a difficult period as a result of a general slackening of world demand for semiconductors. It has already reduced its 2,800 workforce in the United States by three per cent, and is closing its operations for a few days in both September and December to save money. Part of this general retrenchment involves phasing out several products, including the bubble memory devices, which achieved sales of only one million dollars last year.

Bubble memories work by storing information in the form of small polarized regions created in magnetic film. These appear as "bubbles" of magnetism, and can be forced along tracks by a pulsating magnetic field.

They have several advantages over more conventional memory systems, one being a potentially large capacity to store information. Intel announced in 1979, for example, a bubble memory capable of storing one million bits of information, while semiconductor manufacturers are still struggling to produce reliable chips capable of storing 64,000 bits.

Bubble memories are considerably faster than floppy disk storage. They are more reliable than magnetic tapes and disk storage since they have no moving parts. And unlike silicon chips, they do not lose stored data during power failures.

Against these advantages, however, several factors have handicapped bubble memories in the market. For a start they tend to be considerably slower than silicon chips. And while research and development into bubble memories has been aimed at improving their performance, the price of competing systems has been falling at such a dramatic rate that the advantages of bubbles have been insufficient to achieve the volume sales necessary to reduce prices.

While the price of floppy disk storage fell from about 0.04 to 0.01 of a cent for storing a single bit of information between 1977 and 1981, and random access memories from 0.08 to 0.02 over the same period, the cost of bubble memories remains an order of magnitude higher.

Compounding the price disadvantage have been apparent marketing failures by several of the companies. Both Rockwell and Texas Instruments are said to have concentrated excessively on refining the technology of bubble memories while neglecting the ancillary equipment necessary to make them as easy to use as conventional memory systems.

Supporters of bubble memory technology now admit that they are unlikely to overtake the market position of either disks or silicon chips. Where they are more likely to find a niche is in specialized applications where they can capitalize on their particular advantages, for example in defence or factory equipment operating in particularly harsh conditions.

Another potential application is in telecommunications. However, even Bell Laboratories, which pioneered bubble memories in this field admits to frustration in persuading its parent body, American Telephone and Telegraph, to replace disks with bubbles in telephone switching systems. Initial projections that this would take place in the early 1980s have now been put back "two or three years".

David Dickson

New antibiotics

Combined attack

If the hopes of the British pharmaceutical company Beecham are realized, a quiet revolution in antibiotic treatment began yesterday (Wednesday 16 September) with the launch of Augmentin on the UK market.

Although in no sense approaching dangerous levels, there is a steady increase in the proportion of pathogenic bacteria possessing resistance to the widely used beta-lactam (penicillin and cephalosporin) antibiotics. Already several drug companies have responded by introducing so-called "third generation" cephalosporin derivatives (such as cefotaxime (Claforan) from Roussel and moxalactam (Moxam) from Shionogi and Eli Lilly) which are less susceptible to inactivation by the beta-lactamase enzyme present in resistant bacteria. These new cephalosporins, which have a much wider spectrum of activity than the "older" beta-lactam drugs, are being touted as an alternative to combinations of "old" penicillins or cephalosporins with aminoglycosides or some other antibacterials such as chloramphenicol or metronidazole for treating or preventing surgical infections.

Where the new cephalosporins fall down, however, is in the route of administration — they must be given by injection and not orally. In general practice, and for patients not hospitalized, oral administration is much preferred. Because Augmentin is claimed to have at least as broad a spectrum as the new cephalosporins, and can be taken orally, there is already great interest in the drug.

Augmentin is a novel combination; it contains a well established semisynthetic penicillin (amoxicillin) together with clavulanic acid which has no antibacterial activity of its own, but blocks the activity of the beta-lactamase enzyme in bacteria resistant to conventional penicillin therapy. Now this combination has been cleared for sale in the United Kingdom, and the Committee on the Safety of Medicines has now permitted Beecham to recommend its use in most of the infections for which the company had requested clearance.

The attraction of a broad-spectrum antibiotic is that it stands a good chance of being effective when used as an initial treatment, before there has been time to go through the process of identifying the

infecting organisms. And this, of course, is usually the case in general practice. Augmentin is indicated for virtually any bacterial infection provided the bacteria are sensitive, and that means in about 95 per cent of cases. Comparable figures for resistance to cotrimoxazole and ampicillin or amoxicillin are around 84 per cent and 72 per cent respectively.

Applications for Augmentin to be allowed onto other markets, including Europe and the United States, are in the pipeline. The next step after that is to make a grade of Augmentin suitable for use in children (the present one is not), and then perhaps injectable forms will be introduced to compete with the third generation cephalosporins for treating surgical infections.

Charles Wenz

California's Medflies

Who to blame?

Washington

Now that the State of California seems to have brought its recent outbreak of Mediterranean fruit fly under control through intense spraying with the pesticide malathion, a sharp dispute has broken out over what — and more importantly, who — was responsible for an outbreak that at one time seemed to threaten not only California's multi-billion dollar fruit industry, but also the political career of Governor Edmund G. (Jerry) Brown.

Last week, officials of the state's Department of Food and Agriculture issued a report laying the blame squarely at the feet of the federal government. According to department head Richard E. Rominger, the outbreak would have been contained much earlier and without the need for aerial spraying had it not been for the accidental release of over 50,000 supposedly sterile flies by the US Department of Agriculture (USDA), some of which were later shown to be fertile.

However, USDA is refusing to be identified as a principal culprit. A representative of the agency said in Washington last week that although some non-sterile flies were "probably" released by mistake, it was impossible to tell whether this was the main cause of the outbreak since only two fertile flies were discovered among 1,300 dissected, the rest having been found to be sterile.

USDA is also pointing to an apparent laxity in quality controls supposed to check on the sterility of the 8,000 million irradiated flies imported from Peru in an attempt to apply biological pest control methods, for which the state was jointly responsible. USDA protocols suggest that several dozen flies from each batch should be checked for sterility. But in the rush to control the outbreak "the program personnel from both the federal and state agencies decided that only one insect should be checked (from each bag) so the state is really equally to blame".

California's fight against the so-called Medfly, a battle which has been fought several times over the past few years, since the fly is endemic to Central America and is thought to be brought over the border in fruit and vegetables, found itself in the national headlines in July when traditional ground-based control methods, including selective spraying, biological pest control and stripping trees of fruit, appeared to be failing to stop the latest infestation.

Initially Governor Brown refused to allow aerial spraying by malathion because of the general environmental damage he thought this would cause. However, he changed his mind when the federal government threatened to place all Californian produce in quarantine unless the spraying was carried out. After some technical hitches, this began on 20 July.

So far, the massive spraying seems to have been successful in containing the infestation. Although Medfly larvae are still being found at the edges of the sprayed areas — as well as occasionally in other locations to which they are thought to have been transported by humans — none has been found in places where the spraying has been concentrated.

But if the biological battle has been won, at least temporarily, the legal and political battle may be only just beginning. Governor Brown, who draws much of his support from the environmentalist community which opposed the aerial spraying, has been widely criticized by local farmers for not acting sooner. Brown hopes to run for the US Senate when his term as governor expires, a factor said to be playing a large part in his aides' insistence that the federal government should be blamed for the Medfly outbreak, and should therefore help to cover the costs of controlling it.

However, it may prove extremely difficult to determine who is to blame. In addition to the problems with the initial quality control, there is evidence that laboratory technicians who have examined over 100,000 flies have had difficulty in distinguishing the irradiated ones — marked with a yellow dye — from local flies onto which some of the dye has brushed off.

David Dickson

The Medfly is not an aphid, as described in a recent leading article (see *Nature* 27 August, p.786), but *ceratitis capitata*, a dipteran — Editor, *Nature*.

Research in Romania Time for rethink

Scientific research and development work in Romania is being hampered by professional "discoverers" — people who put forward artificial suggestions and "pad" each other's research — according to a recent round-table discussion organized by *Era Socialista*, the theoretical journal of the Romanian Communist Party. At the same time, technologists and

engineers genuinely contributing to industrial research are overburdened with routine paperwork, and some industrial managers try to keep scientists out of their plants on the grounds that they interfere with production. Pay for research staff is too low, and the bonuses payable to those who have put forward a practical new idea often arrive three or four years late. Moreover, no bonuses are paid out on a negative result — and due recognition for negative results is of particular significance in Romania which now has a policy of "reattestation", the conferring of academic degrees not for life but on a renewable basis, depending on recent professional performance.

The round table was part of the new Romanian drive for research and development. During the past 10 years, the research workforce has expanded from 37,000 to 200,000. Since 1976, all research has had to be directly linked with specific needs of the national economy, virtually eliminating the concept of fundamental and academic research. Yet press reviews of the 1976–80 five-year plan indicate that the 12,106 new patents granted during that period earned the national economy some 8,700 million lei, barely enough to offset the state's research and development expenditure during that period.

During the current five years, research expenditure is scheduled to increase sharply. For example, the Minister of Mines, Oil and Geology, Virgil Trofim, told the *Era Socialista* round table that he had been allotted an extra 13,000 million lei for geological research alone — a figure approximately double the total allocation for all scientific research under the previous plan. Last week a decree of the State Council "released" Mr Trofim to take up "other assignments", and divided his former Ministry into three, in terms which suggested that the research funding will be increased even further.

At the end of last year, summing up the research and development shortfall under the 1976–80 plan, President Nicolae Ceausescu criticized in particular the failure to achieve set targets for "technical progress" and failures, or long delays, in introducing many of the results obtained into industry. (Such criticisms, incidentally, must cause the president some conflict between his public and domestic personae. In 1979, when the National Council of Science and Technology was reorganized to strengthen central party and government control, his wife, Dr Elena Ceausescu, was made its head, while in April 1980, two of their children, Valentin and Zoe, became council members.)

In June this year, Dr Ion Ursu, First Vice-Chairman of the National Council of Science and Technology announced a crash research and development programme to achieve over 3,000 main "targets" of national importance, 2,600 of which are to be introduced in industry before the end of the present five year plan.

Vera Rich

Pearls before swine

Bangalore

India and China are showing renewed interest in biogas production as the answer to the energy crisis "down on the farm".

The concept is simple — animal and vegetable waste are converted into methane gas and thence to workable energy. The practical problems of constructing and maintaining suitable digesters and the costs of steel and concrete fittings have, however, proved beyond the capabilities of many third-world farmers. In the Chinese province of Sichuan, for example, more than half of the 7 million digesters have fallen into disuse.

The situation may be eased with the development by the Union Industrial Research laboratory near the Taiwan capital Taipei of an almost indestructible, non-corrosive biogas digester working on pig excreta. The plant is made up of a tough pliant plastic sheet derived from blending red mud with polyvinyl chloride. It is resistant to all types of acids and alkalis, can withstand deflation and inflation 6,000 times per year and is half the cost of conventional digesters used in China. In Taiwan a unit big enough to process the manure of 30 pigs costs around \$225. An adult pig weighing 90 kg excretes about 8 litres of body waste per day; this converts to a third of a cubic metre of gas with a quarter of the heating value of petroleum-derived bottled gas.

Cow dung is the major ingredient for biogas plants in India. Here the most widely used plant, based on a design evolved by the Khadi and Village Commission, consists of an underground cement pit covered by a metal gas holder which had to come from an urban workshop. However, a drumless plant has been designed by a governmental research agency in Uttar Pradesh, which has many similarities with the Chinese biogas plants. It is also half the cost of the conventional plant and can be easily handled and maintained.

It has been estimated that India's cow dung potential could run 18.75 million family-size and 5.6 million community-size biogas plants with a daily capacity of 1.7 million and 142 million cubic metres of gas respectively. The target of 1 million family-size biogas plants, under India's sixth five-year plan, for which \$55 million has been set aside, can only utilize 5 per cent of the potential capacity.

Various new feedstocks are being tested in experimental plants: the Punjab Agricultural University is using leafy material such as tree leaves, crop residues, fruit peel, chopped paddy and wheat straw; and the Indian Agricultural Research Institute in New Delhi is experimenting with both animal and plant wastes.

Radhakrishna Rao

CORRESPONDENCE

Not so civil

SIR — Your comments (*Nature* 292, 396; 1981, and 287, 264; 1980) on the recent *Review of the Scientific Civil Service (1980)* do not mention a major impediment discouraging scientists in the civil service from moving to administrative work. This is the opportunity post system. Open opportunity posts may be filled by suitable staff from any group or class in the service, and an increase in the number of such posts is recommended in the *Review*. However, on appointment to an opportunity post the successful candidate is paid the salary appropriate to his class of origin. In other words there are several rates for the same job. This is obviously unjust. It discriminates against scientists because in the civil service members of the administrative class are paid much more than their equivalents in scientific grades. For example the first secretary (science and technology) in the Paris embassy (an opportunity post at principal "level") may be paid over £1,400 p.a. extra if he happens to be from the civil service rather than the scientific class.

If different salaries were paid for a particular post according to the sex or colour of the incumbent it would be illegal. An equally artificial form of discrimination against scientists moving to administrative work in the civil service is likely to ensure that such moves continue to be few.

C.E. DYTE

Datchet, Slough, UK

Risk of radon

SIR — R.D. Evans *et al.* (*Nature* 12 March, p.98) considered the individual lung cancer risk to the general population from indoor exposure to radon-222 and concluded that a current upper estimate is 10^{-4} per lifetime per working level month. The authorship was remarkable in consisting of six widely recognized senior experts from four countries, and the estimate made was equally noteworthy — an agreement on a risk that was reasoned from cited data that vary over a factor of nearly 50.

It is encouraging that experts can agree, but it is also potentially intimidating. For this reason it is vital to recognize the extent of the continuing uncertainty, and regard the recent estimate only as a starting point rather than a known value. Any number that is selected from within a possible range of 50 times can scarcely be regarded as an upper limit.

Although Evans *et al.* described many of the uncertainties in their estimate, a further potentially major consideration was not included. The general background effect of radon was assessed by identifying lung cancer rates before cigarette smoking became popular. Those rates are, however, not relevant if smoking increases the risk more than additively in a population where most people are exposed to tobacco smoke. Such would be the case if tobacco smoke is a promoter — an agent which accelerates the development of lung cancer but does not of itself nucleate cancer. Data from studies on rats¹ in fact show such an effect — increased lung cancer from exposure to both smoke and radon but no lung cancer from smoke alone. And human data from uranium miners² have

shown a similar enhancement from both smoking and radon. No human data are available where radon is totally absent. Our two references are by no means conclusive, but they make it evident that there is a significant possibility that the appropriate background level of radon-related lung cancer is considerably higher than the value assumed by Evans *et al.*

ROBERT L. FLEISCHER

General Electric Research & Development Center, New York, USA

1. Chameaud, J., Perraud, R., Chretien, J., Massen, R. & Lafuma, J. *Symp. Late Biological Effects of Ionizing Radiation*, 429–436 (IAEA, Vienna, 1978).
2. Lundin, F.E. Jr, Lloyd, J.W., Smith, E.M., Archer, V.E. & Holaday, D.A. *Health Phys.* 16, 571–578 (1969).

SIR — There is some uncertainty, of course, in our estimate of risk, but there is more uncertainty in the proposition that it may be considerably higher.

Authors have put forward various estimates of the risk coefficient for lung cancer in uranium miners, but a value in the range $1-5 \times 10^{-4}$ per working level month commands widest support. It derives from miners who smoked. We argued that a rounded value of 10^{-4} per working level month would be an appropriate upper bound for exposure to radon decay products in the home.

Such a value is seen to account for virtually all the lung cancers in a general population with relatively little smoking. Although one cannot separate the numerical influence of radiation and cigarettes, we are encouraged by the epidemiological circumstances and dosimetric considerations to believe that we are not seriously underestimating the risk.

M.C. O'RIORDAN

National Radiological Protection Board, Oxon, UK

DR O'RIORDAN, a colleague of the late Dr A.S. Maclean, has replied on behalf of Evans *et al.*

Antibiotic practice

SIR — The arguments in your leading article "Saving antibiotics from themselves" (*Nature* 20 August, p.661) are very persuasive. However, in general practice at least the fears of widespread antibiotic resistance have not been realized. In more than thirty years of clinical experience in general practice of both myself and my colleagues, no loss of antibiotic efficacy has been observed. This is not to say that antibiotic resistance is not a problem in the hospital environment.

Antibiotic resistance may be advantageous to the organism in the host but when present in the general environment it might be selected against by measures not yet understood.

Most physicians would agree that a tenfold reduction in the prescribing of antibiotics would probably have negligible effect on the health of patients in advanced countries. However, such a reduction would encounter considerable resistance from the patients themselves and impose delays in the commencement of treatment when it is genuinely required, and additional burdens on bacteriological departments.

Before such measures are undertaken it would be wise to ensure that the theoretical dangers of antibiotic resistance are in fact realized in practice and that organisms possessing antibiotic resistance are not at some

kind of disadvantage when circulating in the general environment which keeps their numbers down to a level at which their dangers remain insignificant.

BRENNIG JAMES

Marlow, Bucks, UK

Fully blind review

SIR — To make the present review system a valuable screening procedure it is necessary that the authors be anonymous just as the referees are. I agree with Dr Gaylarde (*Nature* 30 July, p.402) that the pages containing the author's name, address and also the acknowledgements should be retained in the editorial office when a manuscript is forwarded to referees. However, I am afraid that a referee can still know the names and addresses of the authors if the authors choose to quote their own work in the first person, for example, as follows: "We have previously shown (Jones *et al.* 1980) that . . ." To ensure double-blind evaluation of a given piece of work, I suggest that the authors quote even their own work in the third person, that is, as follows: "Jones *et al.* have previously shown (1980) that . . ." After double-blind evaluation of a manuscript, the authors and the referees, if desired, can be made known to each other.

B. SESHU

Department of Immunology and Microbiology, Wayne State University School of Medicine, Detroit, Michigan, USA

Tamuz bombing

SIR — Your editorial on Israeli bombing was sanctimonious and unrealistic (*Nature* 18 June, p.523). Don't you realize that Israel and Iraq are at war? That Iraq refused to sign a peace agreement?

The only reason Iraq is not bombing Israel now, or ever since the 1967 war, is because Iraq cannot do it — its planes will be destroyed by Israel. So instead, Iraq sends armaments and soldiers to the Palestinians in Lebanon to fight the Israelis. The Iraqis themselves picked an easier target in Iran, or so it seemed at the beginning of their war.

You can't understand what it means to be at war with a country? How the Israelis have to fight enemies in Lebanon, Libya, Iraq, Syria, Jordan? You don't understand what this means?

The headlines in tonight's newspaper, 24 June 1981, read "Iraq's Hussein demands Arabs be given A-bomb". Iraq was building an A-bomb for all the Arabs, not only itself. "If you are looking for evidence, this is it." "The only thing the Iraqis want is an atomic bomb to use against Israel."

Also in *Chem. Eng. News* of 15 June, "The Iraqi reactor could have produced enough plutonium each year for about one bomb".

The "Non-Proliferation Treaty" is a farce. Countries at war do not observe such niceties. As I stated: your editorial is sanctimonious — making a pretence of righteousness and reality. Your readers should expect better from a science periodical.

A.B. CARPENTER

Stoller Research Co., Santa Cruz, California, USA

Agricultural production and malaria resurgence in Central America and India

Georganne Chapin* & Robert Wasserstrom†

AMONG the inhabitants of Asia, Latin America and tropical Africa malaria remains a major cause for alarm. Yet only a few years ago, health officials in a dozen developing countries (capitalizing on the discoveries of British parasitologist Ronald Ross half a century earlier) pointed triumphantly at their efforts to eradicate entirely this mosquito-borne scourge¹⁻⁵. Following World Health Organization (WHO) guidelines, for example, Indian authorities instituted a programme of medical treatment and pesticide application in 1952 which within a single decade reduced the number of cases from over 100 million to 50,000 (ref. 6). Ten years later, using the same methods, health workers in Sri Lanka cut the annual incidence of malaria from three million cases to fewer than 25.

By 1970, however, it had become clear

which mosquitoes of the *Anopheles* subfamily transmit *Plasmodium* parasites to human beings. After reproducing in astronomical numbers, these parasites transform their human hosts into a reservoir of illness which may be spread to uninfected individuals.

In 1907, Dr William Gorgas, an American Army surgeon in the Panama Canal Zone, set out to break this cycle by draining swamps, emptying, covering or oiling pools of standing water and screening human habitations. Although he was unable to eradicate the disease completely, within two years the death rate from malaria among canal company employees had fallen to 8.86 per thousand — a decrease of 80 per cent. More or less simultaneously an Italian physician, Dr Angelo Celli, noticed that in southern Europe the disease tended to attack people

for a short time on the inside walls of human houses after biting their victims. By spraying walls with DDT, WHO officials reasoned they could reduce the vector population to manageable levels. Simultaneously, they proposed to treat everyone affected by the disease with chloroquine or other anti-*Plasmodium* drugs so as to destroy the reservoir of disease. If these conditions could be maintained for three or four years, they calculated, malaria transmission might be broken forever. Moreover, projects of this sort could be carried out without altering political arrangements or patterns of land tenure.

Complications

Despite the simplicity of the plan, WHO officials were aware that a worldwide campaign to eradicate malaria would face nearly insurmountable obstacles. In many regions, control programmes were virtually non-existent, and even those that did exist suffered from a lack of funds, technical expertise and administrative efficiency. But there was another obstacle which experts at WHO were more reluctant to engage: as early as 1953, they obtained conclusive evidence that *Anopheles* mosquitoes, like many insect pests, sooner or later became resistant to DDT and other pesticides. Within a few years, in fact, such resistance had been reported in Greece and Italy (where insecticides were used both in public health and in agriculture) as well as in the Lebanon, Iran, Saudi Arabia and Nigeria. In some cases a single application was sufficient to reduce mortality (that is, increase resistance) among mosquitoes by 80 per cent¹⁸. Accordingly, WHO malariologists urged their local counterparts to conduct "time-limited" spraying operations — to complete the "attack" phase as quickly as possible.

Thus, anti-malaria teams were directed to treat the interior walls of all human habitations and shelters within the target zone on a regular schedule — a gargantuan task under the best of circumstances. Meanwhile, by organizing an elaborate

A few years ago the battle against malaria seemed to have been won. However now, despite vigorous anti-malaria campaigns, the disease has made a comeback. The resurgence of malaria in Central America and India seems to have been paralleled by intensified agriculture in these countries and the associated increased use of pesticides.

that malaria eradication had run into severe difficulties. Instead of dwindling to insignificance, the number of infected individuals rose again to distressing proportions. In India, which had served as a showplace for WHO policies, five million people were soon infected; in Sri Lanka, two million people became sick again almost overnight; and in Central America infection rates grew to previously unknown levels⁷. Moreover, unlike earlier outbreaks, this new plague was often carried by mosquitoes which had become resistant to pesticides like DDT and dieldrin and could not be controlled by conventional means⁸⁻¹⁵. The origins of this major ecological disaster must be sought as much in the unwitting actions of international organizations as in hapless nature.

A seeming success

It is worth noting that early programmes to contain parasitic diseases — primarily malaria and yellow fever — achieved remarkable success without recourse to sophisticated technologies¹⁶. Efforts to overcome malaria before the Second World War concentrated on the ways in

who were either poor or landless, particularly those who worked as seasonal labourers on certain large farms. He reasoned that transmission depended to a considerable degree on the flow of fresh human blood into malarious zones just as the *Anopheles* population began its annual explosion. Not only sanitary measures and medical treatment, then, but land reform and other social programmes might play a substantial part in conquering the disease.

By the end of the Second World War, however, it seemed that a technology had finally been devised that could eliminate malaria in much of the non-Western world: chemical pesticides. Insecticides like DDT and dieldrin were not only cheap and easy to use, but they also remained active for several months after each application¹⁷. For this reason, they appeared to be ideally suited for the task of killing *Anopheles* mosquitoes, which characteristically rested

Table 1 Prevalence of malaria in Central America, 1965-1977

	No. of cases (thousands)													
Country	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	
Costa Rica					0.6	0.3	0.3	0.2	0.2				0.2	
El Salvador	34.2	68.6	83.0	35.8	25.3	45.4	46.8	38.3	35.1	67.0	83.1	83.3	32.2	
Guatemala	14.3	22.0	21.2	11.0	10.6	10.9	8.3	9.3	1.0		1.0	9.6		
Honduras	6.9	17.1	16.1	15.7	29.6	34.5	48.4	18.6	8.8	7.5	30.3	48.8	39.4	
Nicaragua	8.3	15.6		7.1	16.0	28.5	23.3	9.6	4.2	12.2	24.7	26.2	11.6	

Source: WHO, *World Health Statistics Annual*, Geneva, 1966-79.

*Division of Sociomedical Sciences, School of Public Health, Columbia University, New York, New York 10032, USA.

†Department of Anthropology, and Division of Sociomedical Sciences, Columbia University, New York, New York 10027, USA.

system of regional laboratories and clinics, public health officials were supposed to administer chemotherapy and monitor the campaign's progress. In areas where these tactics proved to be successful, where the number of active cases diminished to zero, attack gave way to consolidation. And if no new illness occurred during the following three years, consolidation was in turn replaced by maintenance, the constant vigil against a recurrence of infection. This strategy was adopted in 1954 by the Pan American Health Organization (PAHO) and subsequently by the entire international community.

Initially, at least, it seemed that WHO's campaign enjoyed almost unmitigated success. In India, after ten years of struggle against malaria (1961), only 50,000 cases of the disease were uncovered by government inspectors and a number of regions had passed from attack to consolidation or maintenance. Similar triumphs were registered in Pakistan, Sri Lanka, Paraguay, Venezuela, Mexico and Central America. In ten other countries, *Plasmodium* infection was completely overcome.

Malaria re-emerges

Within a short time, however, the campaign began to falter. Between 1961 and 1966, disease rates in India increased threefold; by 1970, half a million people caught malaria each year — many in areas where health authorities had recently scored impressive victories. Much the same course of events took place in Sri Lanka, which in 1968 experienced an epidemic that left 1.5 million people stricken. On the other side of the world, in El Salvador, Nicaragua and Honduras (where anti-malaria measures began in the late 1960s), the incidence of disease in 1975 was three times greater than a decade earlier, before the programme had started (Table 1). As a result, eradication projects which had reached consolidation frequently reverted to the attack phase — or even entered the newly defined stage of "permanent attack". Even so, it soon became clear that there was a major resurgence of malaria in India and Central America that existing administrative and technological methods could do little to prevent¹⁹⁻²⁴. The question which malariologists in these areas then asked themselves was quite simply, "What has gone wrong?"

In fact, as early as 1962 a number of specialists had expressed their reservations about the WHO campaign and its chances of success. Among other things, they pointed out that as infection rates dropped during the attack phase, hard-pressed governments often diverted critical resources from anti-malaria activities to other essential projects^{25,26}. As a result, many infected people were not detected by surveillance systems, which themselves broke down under poor management and supervision. Even more ominously, however, resistance to DDT and dieldrin

Size of holding (hectares)	No. of holdings (thousands)	% Of rural properties	Area (millions of hectares)	% Of total area
< 0.5	23,178	32.9	5.4	3.4
0.5- 1.0	12,504	17.7	9.1	5.6
1.0- 1.9	13,432	19.1	19.3	11.9
2.0- 2.9	6,722	9.5	16.4	10.0
3.0- 3.9	3,959	5.6	13.6	8.4
4.0- 4.9	2,684	3.8	11.9	7.4
5.0- 9.9	5,248	7.4	36.3	22.4
10.0-19.9	2,135	3.0	28.5	27.6
20.0-29.9	401	0.6	9.3	5.8
30.0-39.9	120	0.2	4.2	2.6
40.0-49.9	45	0.1	2.1	1.3
>50	65	0.1	6.0	3.1
Total	70,493	100.0	162.1	100.0

Source: Ministry of Planning, Government of India, *Statistical Abstract, India, 1977*, New Delhi, 1978.

had reached alarming proportions among *Anopheles* mosquitoes — just as WHO officials had originally feared^{27,28}.

The case of El Salvador is illuminating. In 1958 a group of entomologists reported that the local vector, *Anopheles albimanus*, had lost its susceptibility to all major organochlorine compounds and was proliferating rapidly along the Pacific coast^{29,30}. Four years later, researchers in southern Mexico encountered the same problem, which forced them to admit that the disease had not been eradicated in several areas³¹. In India, widespread tolerance to organochlorines was discovered among two important vectors, *Anopheles culifacies* and *Anopheles fluviatilis*, particularly in regions which had recently shifted to high-yielding forms of agricultural production³².

In such places, effective control might be regained only by using insecticides which cost four, five or even ten times as much as common toxins — a burden which few governments were willing to bear³³. Yet

even measures of this kind might serve at best only as temporary expedients: vectors which became resistant to one compound frequently enjoyed mysterious immunity to other unrelated poisons, and in any case it was only a matter of time before natural selection favoured those insects which could withstand a broad spectrum of chemical agents³⁴⁻³⁶. Faced with these problems, in 1973 WHO officials reluctantly transformed the Malaria Eradication Division into the Division of Malaria and other Parasitic Disease^{37,38}.

Agricultural expansion

Against this background, it is ironic that commercial agriculture often expanded in precisely those regions recently cleared of malaria. As the danger of illness subsided, many landowners, stimulated by the high prices of such commodities as cotton, rice and tobacco, reduced their production of other crops and bought up more land³⁹. The result of this development was both to augment the number of poor and landless rural workers and to concentrate agricultural resources in progressively fewer hands⁴⁰. In 1960, for example, the wealthiest five per cent of India's non-urban population owned nearly one-third of its croplands; ten years later, making use of lavish credit facilities and development funds, they had actually increased their holdings by ten per cent (Table 2)⁴¹.

Naturally, such proprietors also claimed a disproportionate share of the country's agricultural income: between 1960 and 1970, in fact, the richest 20 per cent of these people (like their counterparts in Pakistan and Bangladesh) collected nearly half of such earnings⁴². Similarly, in Central

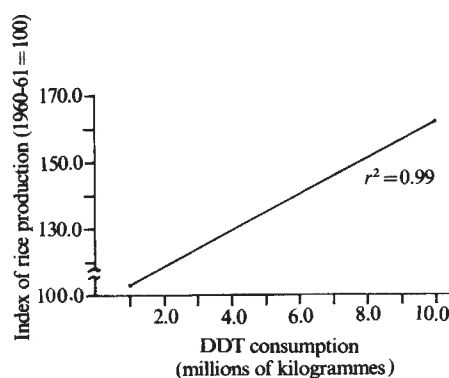


Fig.1 Effect of DDT use on rice production in India, 1970-77

America, where cotton became king in the 1960s, both land and credit remained almost completely inaccessible to most small growers. In El Salvador, one per cent of the rural population owned half of the country's land (including all properties over 50 hectares) and received 30 per cent of its income^{43,44}. As in India, such prosperity was achieved only at the expense of considerable inequity: one-third of the area's peasants possessed less than one hectare and one-quarter of them owned no land at all. Clearly, the expansion of commercial agriculture in these regions exacerbated such difficulties and created an entirely new kind of agroecosystem.

Increased pesticide use

In these circumstances, it is not surprising that the death rate from infectious and parasitic diseases in Central America has remained extremely high and that the incidence of malaria has generally increased — despite an impressive diminution in the late 1960s and early 1970s⁴⁵. The relationship between fibre production and the recrudescence of malaria has been clearly established in a study by the United Nations Environmental Programme and the Instituto Centroamericano de Investigación y Tecnología Industrial (ICAITI)⁴⁶. To combat cotton pests and to raise yields, planters in Guatemala, Nicaragua and El Salvador have not only expanded their acreage, but since 1970 they have also applied heavier concentrations of pesticides. Whereas a decade ago, fields were sprayed only eight or nine times each season, they must now be fumigated on as many as 50 occasions. Consequently, the amount of pesticide which enters the local ecosystem has expanded at an increasing rate.

In 1971, for example, farmers in El Salvador sprayed 58.4 kilos on each hectare of cotton; three years later, applications had reached 70.0 kg per hectare. As a result, DDT consumption in El Salvador increased threefold between 1970 and 1977 — from 555,200 kg to 1.6 million kg. Similar circumstances prevailed in Nicaragua, where DDT imports rose from 29,000 kg in 1974 to 521,600 kg in 1976.

Naturally, the importation of pesticides on this scale could be accomplished only as long as cotton revenues offset the rising cost of toxins. Fearful that unstable prices and soaring expenses might soon cut their earnings and anxious to maximize the returns on their investments, many growers have attempted to achieve total control of insect parasites — an obsession which has only enhanced their reliance on expensive chemicals. Correlating the use of DDT in El Salvador with renewed malaria transmission, it can be estimated that at current rates each kilo of insecticide added to the environment will generate 105 new cases of malaria⁴⁷⁻⁴⁹.

Ironically, *Anopheles* resistance in India has developed even in areas where cotton is

grown on relatively small plots of land or where food grains — primarily rice — still predominate. In Tamil Nadu, for example, most farmers (77.5 per cent) own less than one hectare; few possess more than two⁵⁰. Even so, these men (and their counterparts in Maharashtra and Gujarat) produce nearly one-third of the country's cotton and an impressive share of its rice^{51,52}. By 1968, too, the incidence of malaria in this region had decreased to insignificance — indeed, in Tamil Nadu, the disease was

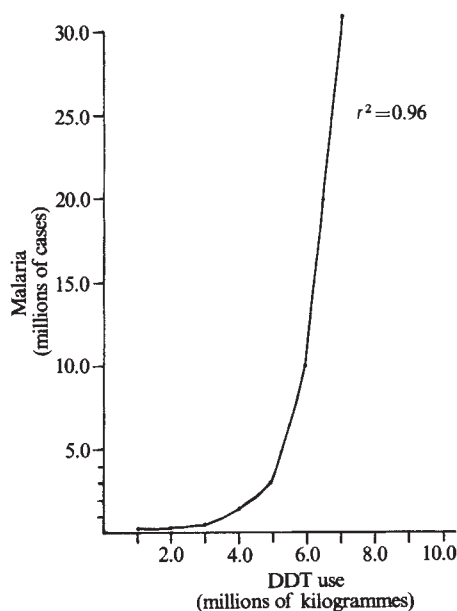


Fig. 2 Relationship between DDT use and malaria incidence in India, 1969–77

largely confined to urban areas such as Madras⁵³⁻⁵⁶. Within a few years, however, public health officials throughout southern India reported that mosquitoes of both the *Anopheles* and *Aedes* subfamilies (the latter transmit yellow fever) had become resistant to a wide variety of chemicals including DDT, BHC, malathion (an organophosphate) and propoxur (a carbamate)⁵⁷⁻⁶¹.

Resistance to pesticides

Indisputable evidence is not yet available but it appears that resistance began to occur with the introduction of green revolution technology — particularly of high-yielding varieties (HYV) of rice (Fig. 1). According to entomologists at the Vector Control Research Centre in Pondicherry, “the major changes that have been taking place in the area are in the tremendous increase of acreage under cultivation, the near total replacement of organic manure by chemical fertilizers and the extensive use of insecticides for paddy and other crops”⁶². It should be noted that the new strains of rice have been adopted primarily by wealthy landowners and have proved to be especially susceptible to insect infestation⁶³⁻⁶⁶.

Recent studies in Tamil Nadu indicate that families which own two hectares or more (7.2 per cent of the total number)

alone possess the means to purchase HYV seeds, fertilizers and pesticides⁶⁷⁻⁶⁹. As Figs 2 and 3 suggest, such farmers have responded to increased infestation by applying heavy doses of DDT, BHC and dieldrin — a procedure which is closely related to the recent explosion of malaria in the region. Significantly, as these growers have switched from DDT to more sophisticated chemicals, traditional vectors have been replaced by rarer species which in turn show a diminished sensitivity to such poisons⁷⁰. Little wonder, therefore, that as early as 1972 the *Indian Journal of Public Health* warned that “the most serious threat to public health . . . is the uncontrolled use of pesticides for agronomic practice”⁷¹.

Deadly link

There seems to be a three-stage relationship between the evolution of cotton agroecosystems and the spread of malaria. During the first stage, eradication programmes are more or less effective and often permit farmers to exploit previously infected areas. An example is to be found in eastern Paraguay, where landless peasants from so-called overpopulated regions have been encouraged to clear and colonize “uninhabited” jungle in which malaria is endemic⁷². The Paraguayan government has made great efforts to eliminate the *Plasmodium* parasite — as a stimulus to both immigration and productivity. Among its primary objectives, the production of cotton and tobacco (also a heavy user of pesticides) for export occupies a pre-eminent position. By 1978, these two crops provided over 27 per cent of the country's foreign exchange. Having prepared vast expanses of virgin forest for commercial exploitation, however, peasants in many areas have begun to abandon their farms and to look for wage labour elsewhere. As their lands are consolidated into larger holdings, such areas will inevitably be treated with intensive applications of DDT or dieldrin. These lands will therefore probably come to resemble many parts of Central America (stage two), in which the high price of cotton appears to justify augmented doses of pesticides. Malaria transmission will be stimulated most markedly among migrant workers and impoverished peasants. Perhaps in anticipation, PAHO has already commissioned a study in Paraguay, *The Impact of Malaria on Economic Development*⁷³, according to which seasonal increases in *Plasmodium* infection do not interfere with cotton or tobacco cultivation — although they may wreak havoc on food production. Finally, in places like India, Pakistan and Bangladesh (stage three), more and more DDT must be sprayed simply to maintain a fixed yield (Fig. 4). In this case, pesticide addiction and a full-fledged epidemic of malaria have entered their most destructive phase.

So must countries such as India and El Salvador cease to grow cotton and high-yielding food grains? Can foreign currency be obtained only at the expense of widespread malnutrition and disease? It is instructive to examine how such crops are produced in the United States. American farmers who started raising cotton with only small quantities of insecticides found that insect pests reappeared in their fields almost as soon as the fumigators departed⁷⁴⁻⁷⁸. Most often, they responded by applying stronger poisons with greater frequency until they were spraying every two or three days for five months⁷⁹. The side effects had by then become apparent: cattle fodder had to be destroyed because it contained pesticide residues too high to be fed to animals while crops that had never suffered severe infestations were suddenly devastated by previously innocuous insects⁸⁰⁻⁸⁴.

Integrated pest management

In response, entomologists developed what they call integrated pest management systems^{85,86}, the key to which lies in timing insecticide applications so that the crop is protected from predators only at the most vulnerable stages of its growth cycle. As it turns out, cotton buds destroyed by pests regrow throughout the plant's life, so that producers can afford to sustain a high level of insect damage before there is a need to apply pesticides. Simple precautionary measures may also lower their chemical costs: up to 75 per cent of the hibernating boll weevil population may be eliminated by the ploughing under of crop debris after harvest. Thus many growers west of the Mississippi now spray their fields only seven or eight times each season instead of 25 or 30; similar measures have been developed for raising corn, rice and many kinds of fruit⁸⁷.

So why did WHO not urge cotton producing countries to employ integrated management systems that would not

interfere with malaria eradication programmes? A possible answer may perhaps be found in the activities of another international agency, the Food and Agricultural Organization (FAO). Like WHO, FAO was established to provide technical advice and assistance to members of the United Nations. In the case of pesticides, which are manufactured and distributed by a few multinational corporations, FAO's advice might have played a critical role in reducing environmental contamination. Both farmers and extension agents in developing nations must normally rely on pesticide company salesmen for information about how to use agricultural chemicals — much as physicians in Western countries rely upon pharmaceutical companies for information about new drugs. Beginning in 1967, therefore, FAO put together a small working group of experts on integrated pest management which published technical manuals and disseminated other information⁸⁸⁻⁹⁴.

Three years later, it commissioned an American entomologist, Dr Louis Falcon, to develop an integrated system in Nicaragua, a system which achieved remarkable success within a few seasons. Similar programmes were subsequently undertaken in Mexico, Peru and Pakistan⁹⁵. Then, in 1975, FAO delegates met in Rome to consider the question of pesticides in agriculture and public health. Although they recognized that integrated pest management offered a potential solution to many health problems, they recommended that FAO place its emphasis on teaching growers in developing nations how to make more "safe and efficient" use of pesticides^{96,97}.

Whys and wherefores

Why did FAO choose this course of action, which in retrospect does not appear to have been guided by an accurate appreciation of the perils of pesticide addiction? It is important to examine how pesticide manufacturers have influenced the policies of international agencies. As public concern about the effects of toxins like DDT began to grow in the 1960s, these corporations formed a trade association called GIFAP (Groupement International des Associations Nationales de Pesticides) which in turn worked directly with UN technicians through a FAO bureau known as the Industry Cooperative Programme (ICP). By the early 1970s joint FAO-ICP regional seminars had been organized in many parts of the world to promote new and better ways of distributing agricultural pesticides. More important, high-level officials in WHO and FAO, who share the industry's views on many major issues, invited GIFAP to play an active part in agency "consultations" and other internal meetings^{98,99}. In this way, for example, no fewer than 25 corporate representatives lent their expertise to the meeting in Rome

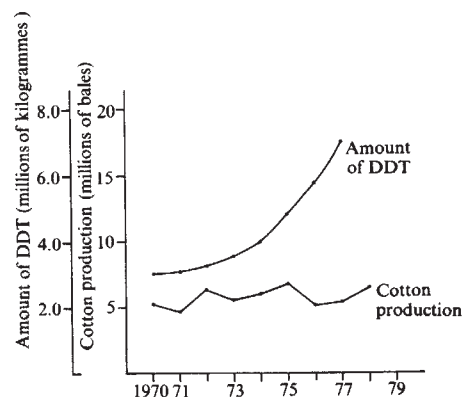


Fig. 4 Cotton production and DDT use in India, 1970-78

on pesticides in agriculture and public health and served on subcommittees responsible for formulating UN policy. Not surprisingly, these subcommittees stressed the need to apply more pesticides in a more effective manner rather than to limit their use or replace them with alternative forms of pest control. And what is more curious, none of these deliberations included representatives of other international constituencies such as environmental groups, labour unions or farmers' organizations. Perhaps for these reasons, in June 1978, the current director general of FAO, Eduard Saoumi, finally expelled ICP from his agency¹⁰⁰.

The problem

In 1976, WHO published a technical report entitled *Resistance of Vectors and Reservoirs of Disease to Pesticides*¹⁰¹. In this report, WHO's Expert Committee on Insecticides declared that "resistance is probably the biggest single obstacle in the struggle against vector-borne diseases and is mainly responsible for preventing successful malaria eradication in many countries . . . Evidence has also accumulated to show conclusively that resistance in many vectors has been caused as a side-effect of agricultural pesticide usage". Accepting this unhappy fate, the report concluded that "vector control was likely to depend on substantial, continued use of pesticides for at least a decade". In these circumstances, it foresaw no alternative but to "encourage commercial firms to continue the search for pest control, especially compounds with a novel mode of action". And yet, as many specialists have pointed out, such compounds are unlikely to resolve this dilemma or to undo the damage which in-bred tolerance has already caused: detoxification appears to rely on physiological processes which are both irreversible and difficult to disrupt. In effect, throughout southern India, the recrudescence of malaria now represents a social cost of growing high-yielding rice — just as elsewhere in India and Central America it represents a social cost of producing cotton.

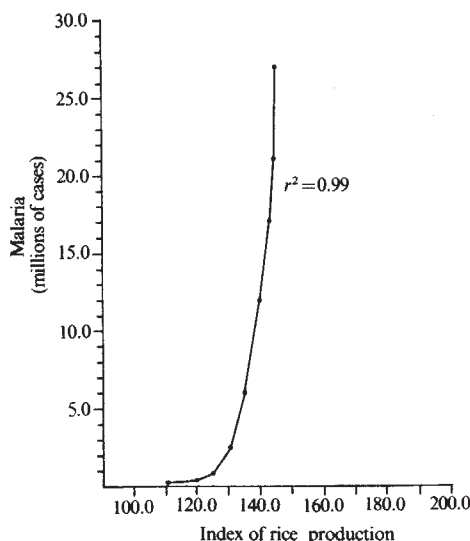


Fig. 3 Relationship between rice production and malaria incidence in India, 1969-77

Recognizing such difficulties, in the same year WHO — joined by the World Bank and the United Nations Development Programme — established its Special Programme of Research and Training in Tropical Diseases (TDR) to coordinate efforts against six ancient scourges of mankind (malaria, schistosomiasis, the filariases, trypanosomiasis, leishmaniasis and leprosy)¹⁰². By June 1980, TDR had allocated over \$2.8 million to 38 separate projects concerned with problems of *Anopheles* resistance and *Plasmodium* transmission. Unfortunately, however, these projects tend to reflect WHO's conviction that progress in the field of malaria research depends almost uniquely on basic breakthroughs in such areas as biology, immunology and the invention of new drugs. So far, only 7.1 per cent of TDR's malaria funds support inquiries into non-chemical means of vector control¹⁰³. Or as Peter Collins has written, the programme draws its models from "many of the ideas and techniques developed by biomedical research into the diseases of the indus-

trialised, affluent world"¹⁰⁴.

A solution?

Must poor countries inevitably choose between more productive forms of agriculture and the uncontrollable spread of parasitic disease, between the irreconcilable claims of "food versus health", as several writers have recently proposed? The evidence, based on a common pattern of social and epidemiological factors, appears to suggest another conclusion. In almost every major case of malaria resurgence, large landowners overuse pesticides on crops like cotton and tobacco which make no substantial contribution to the subsistence requirements of the rural poor. Instead, income from such ventures is generally channelled into luxury consumption at home or speculative investments abroad; only in rare instances does it serve to finance the production of basic foodstuffs or capital equipment.

Even more distressing, much the same situation now prevails in India's main rice-growing regions, where the technology of

HYV rice cultivation has been available principally to wealthy farmers who sell their grain at premium prices. Many specialists now feel that the green revolution in such regions has not significantly improved the lot of most peasants and smallholders¹⁰⁵.

Acknowledging this fact, institutions like the International Rice Research Institute and the International Maize and Wheat Improvement Center have begun to re-evaluate the advantages both of traditional plant varieties and of native cultivation techniques. Perhaps the problem of malaria resurgence is amenable to similar solutions, not by means of technological innovation alone, but also by showing a proper respect for the social and environmental conditions which make resurgence unavoidable. And while more equitable answers are found to the question of poverty and landlessness, surely it makes sense (as FAO's own panel of experts has once again reaffirmed) to employ an integrated system of pest management on such crops as rice and cotton.

1. Brown, A.W.A., Haworth, J. & Zahar, A.R. *J. med. Ent.* **13**, 1-25 (1976).
2. Dutta, H.M. & Dutt, A.K. *Soc. Sci. Med.* **12**, 69-84 (1978).
3. Gabaldón, A. in *Vector Control and the Recrudescence of Vector-Borne Diseases*, 23-24 (Pan American Health Organization, Washington DC, 1972).
4. Lepes, T. *J. trop. Med. Hyg.* **77**, Suppl., 47-53 (1974).
5. WHO tech. Rep. Ser. No. 537 (World Health Organization, Geneva, 1974).
6. Harrison, G. *Mosquitoes, Malaria and Man: A History of the Hostilities Since 1880* (Dutton, New York, 1978).
7. *World Health Statistics Annual* (World Health Organization, Geneva, yearly updates).
8. Ariaratnam, V. & Georgiou, G.P. *Bull. Wild Hlth Org.* **51**, 655-659 (1974).
9. Brader, L. *A. Rev. Ent.* **24**, 225-254 (1979).
10. Breehland, S.G. et al. *Bull. Wild Hlth Org.* **43**, 627-631 (1970).
11. Brown, A.W.A. in *Vector Control and the Recrudescence of Vector-Borne Diseases*, 59-63 (Pan American Health Organization, Washington DC, 1972).
12. Busvine, J.R. *Nature* **273**, 604-607 (1978).
13. Hadjinicolaou, J. & Betzios, B. *Bull. Wild Hlth Org.* **48**, 699-703 (1973).
14. Jeffery, G.M. *Am. J. trop. Med. Hyg.* **25**, 361-271 (1976).
15. WHO tech. Rep. Ser. No. 585 (1976).
16. Harrison, G. *Mosquitoes, Malaria and Man: A History of the Hostilities Since 1880* (Dutton, New York, 1978).
17. McIntire, G. *Environment* **14**, 14-29 (1972).
18. PAHO scient. Publ. No. 238 (1972).
19. Garcia, R. *Environment* **14**, 2-9 (1972).
20. WHO Chron. **32**, 9-17 (1978).
21. Mason, J. *Am. J. trop. Med. Hyg.* **24**, 581-585 (1975).
22. Rachou, R.G., Schinazi, L.A. & Moura Lima, M. *Rev. bras. Malar. Doenc. trop.* **25**, 1-293 (1973).
23. Roy, R.G. et al. *Ind. J. med. Res.* **69**, 53-59 (1979).
24. Roy, R.G. et al. *Ind. J. med. Sci.* **30**, 313-316 (1976).
25. Harrison, G. *Mosquitoes, Malaria and Man: A History of the Hostilities Since 1880* (Dutton, New York, 1978).
26. WHO tech. Rep. Ser. No. 537 (1974).
27. Davidson, G. & Zahar, A.R. *Bull. Wild Hlth Org.* **49**, 475-483 (1973).
28. Brown, A.W.A., Haworth, J. & Zahar, A.R. *J. med. Ent.* **13**, 1-25 (1976).
29. Garcia Martin, G. & Nájera-Morrondo, J. *Boln Of. Sanit. Panam.* (Engl. edn) **6**, 15-23 (1972).
30. Rachou, R.G., Schinazi, L.A. & Moura Lima, M. *Rev. bras. Malar. Doenc. trop.* **25**, 1-293 (1973).
31. de Zulueta, J. & Garrett-Jones, C. *Am. J. trop. Med. Hyg.* **14**, 63-77 (1965).
32. Mojamdar, V.G. *Ind. J. publ. Hlth* **29**, 107-114 (1975).
33. Ghatak, S. & Turner, R.K. *Food Policy* **3**, 136-46 (1978).
34. Benson, R.L. *BioScience* **21**, 1160-1165 (1971).
35. Georgiou, G.P. *A. Rev. Ecol. Syst.* **3**, 133-168 (1972).
36. Plapp, F.W. *A. Rev. Ent.* **21**, 179-197 (1976).
37. Harrison, G. *Mosquitoes, Malaria and Man: A History of the Hostilities Since 1880* (Dutton, New York, 1978).
38. WHO Chron. **32**, 226-230 (1978).
39. *Agricultural Land Settlement* (World Bank, Washington DC, 1978).
40. George, S. *How the Other Half Dies* (Allanheld, New York, 1977).
41. Singh, I. *World Bank Staff Working Pap.* No. 320 (1979).
42. Ojha, P.H. & Bhatt, V.V. *Sankhya* **36**, 163-166 (1974).
43. Dörner, P. & Quirós, R. *J. Latin Am. Stud.* **5**, 217-232 (1973).
44. *Informe General Sobre el Desarrollo Agropecuario y Rural de El Salvador* (World Bank/Interamerican Development Bank, Washington DC, 1976).
45. PAHO scient. Publ. No. 364 (1978).
46. *An Environmental and Economic Study of the Consequences of Pesticide Use in Central America — Cotton Production* (Instituto Centroamericano de Investigación y Tecnología Industrial, Guatemala, 1977).
47. FAO Prod. Yb. **31** (1977).
48. FAO Trade Yb. **31** (1977).
49. *World Health Statistics Annual* (World Health Organization, Geneva, yearly updates).
50. Mencher, J. *Agriculture and Social Structure in Tamil Nadu* (Carolina Academic, Durham, North Carolina, 1978).
51. *Statistical Abstract, India, 1977* (Ministry of Planning, Government of India, Delhi, 1978).
52. *Statistical Outline of India* (Tata Services Ltd, Bombay, 1978).
53. Das, P.K., Reuben, R. & Batra, C.P. *Ind. J. med. Res.* **69**, 403-411 (1979).
54. Dutta, H.M., Dutt, A.K. & Vishnukumari, G. *Soc. Sci. Med.* **13**, 191-194 (1979).
55. Rajagopalan, P.K., Chandras, R.K. & Paniker, K.N. *Ind. J. med. Res.* **69**, 589-597 (1979).
56. Roy, R.G. et al. *Ind. J. med. Res.* **67**, 947-952 (1978).
57. Kaul, H.N. et al. *Ind. J. med. Res.* **64**, 1760-1768 (1976).
58. Mojamdar, V.G. *Ind. J. publ. Hlth* **29**, 107-114 (1975).
59. Rajagopal, R. *Ind. J. med. Res.* **66**, 27-28 (1979).
60. Roy, R.G. et al. *Ind. J. med. Res.* **68**, 744-746 (1978).
61. Pant, C. & Gratz, N. *Outlook Agric.* **10**, 111-115 (1979).
62. Rajagopalan, P.K., Chandras, R.K. & Paniker, K.N. *Ind. J. med. Res.* **69**, 589-597 (1979).
63. Apple, J.L. *BioScience* **22**, 461-463 (1972).
64. Hargrove, T.R. & Cabanilla, V.L. *BioScience* **29**, 731-735 (1979).
65. Smith, R.F. *Bull. ent. Soc. Am.* **18**, 7-14 (1972).
66. Chakravarti, A. *Ann. Ass. Am. Geogr.* **63**, 319-330 (1973).
67. Mencher, J. *Agriculture and Social Structure in Tamil Nadu* (Carolina Academic, Durham, North Carolina, 1978).
68. Harris, J. *Dev. Change* **11**, 33-64 (1980).
69. Chakravarti, A. *Ann. Ass. Am. Geogr.* **63**, 319-330 (1973).
70. Das, P.K., Reuben, R. & Batra, C.P. *Ind. J. med. Res.* **69**, 403-411 (1979).
71. *Ind. J. publ. Hlth* **16**, 105-106 (1972).
72. *Paraguay: Regional Development in Eastern Paraguay* (World Bank, Washington DC, 1978).
73. Conly, D.N. *PAHO scient. Publ.* No. 297 (1975).
74. Bottrell, D.G. & Adkisson, P.L. *A. Rev. Ent.* **22**, 451-481 (1977).
75. Cross, W.H. *A. Rev. Ent.* **18**, 17-46 (1973).
76. Glass, E.H. & Thurston, H.D. *BioScience* **28**, 109-115 (1978).
77. Irving, G.W. Jr *Science* **168**, 1419-1424 (1970).
78. Philogene, B.J.R. *BioScience* **22**, 715-717 (1972).
79. van den Bosch, R. *The Pesticide Conspiracy* (Anchor/Doubleday, New York, 1980).
80. Aldrich, S.A. *BioScience* **22**, 90-95 (1972).
81. Luck, R.F., van den Bosch, R. & Garcia, R. *BioScience* **27**, 606-611 (1977).
82. Weir, D. & Shapiro, M. *Circle of Poison: Pesticides and People in a Hungry World* (Institute for Food and Development Policy, San Francisco, 1980).
83. Pimentel, D. et al. *BioScience* **28**, 772-784 (1978).
84. van den Bosch, R. *Environment* **14**, 18-31 (1972).
85. Bottrell, D.G. & Adkisson, P.L. *A. Rev. Ent.* **22**, 451-481 (1977).
86. Ruesink, W.G. *A. Rev. Ent.* **21**, 27-44 (1978).
87. Kiritani, K. *A. Rev. Ent.* **24**, 279-312 (1979).
88. *Integrated Pest Control* (Food and Agricultural Organisation, Rome, 1977).
89. *Report on FAO Expert Consultation on Pesticides and the Environment* (AGP: 1977/M/7, Food and Agricultural Organisation, Rome, 1977).
90. *Report of the Fifth Session of the FAO Panel of Experts on Integrated Pest Control* (AGP: 1975/M/2, Food and Agricultural Organisation, Rome, 1975).
91. *The Development and Application of Integrated Pest Control in Agriculture: Formulation of a Cooperative Global Programme* (AGP: 1974/M/8, Food and Agricultural Organisation/United Nations Environmental Programme, Rome, 1975).
92. *Impact Monitoring of Agricultural Pesticides* (AGP: 1976/M/4, Food and Agricultural Organisation/United Nations Environmental Programme, 1976).
93. Falcon, L.A. & Smith, R.F. *Guidelines for Integrated Control of Cotton Insect Pests* (FAO/AGPP: MISC/8, Food and Agricultural Organisation, Rome, 1973).
94. *Integrated Pest Control* (AGP: 1979/M/1, Food and Agricultural Organisation/United Nations Environmental Programme, Rome, 1979).
95. Ling, L.L. *A. Rev. Ent.* **19**, 177-196 (1974).
96. *Report of the Ad Hoc Government Consultation on Pesticides in Agriculture and Public Health* (AGP: 1975/M/3, Food and Agricultural Organisation, Rome, 1975).
97. *Impact Monitoring of Residues from the Use of Agricultural Pesticides in Developing Countries* (AGPP: MISC/22, Food and Agricultural Organisation, Rome, 1975).
98. *Report of the Ad Hoc Government Consultation on Pesticides in Agriculture and Public Health* (AGP: 1975/M/3, Food and Agricultural Organisation, Rome, 1975).
99. George, S. *How the Other Half Dies* (Allanheld, New York, 1977).
100. Taylor, W. *Multinational Monitor* **2**, 24-26 (1980).
101. WHO tech. Rep. Ser. No. 585 (1976).
102. *Action Against Tropical Diseases* (United Nations Development Programme/World Bank/World Health Organisation Special Programme for Research & Training in Tropical Diseases, Geneva, 1978).
103. *Fourth Annual Report, 1 July 1979-30 June 1980* (UN Development Programme/World Bank/World Health Organisation Special Programme for Research & Training in Tropical Diseases, Geneva, 1980).
104. Collins, P. *Nature* **269**, 550-551 (1977).
105. Pearse, A. *Seeds of Plenty, Seeds of Want* (Clarendon/Oxford University Press, Oxford, 1980).

NEWS AND VIEWS

Supernovae

from Virginia Trimble

SUPERNOVAE as a class of astronomical events releasing $\geq 10^{49}$ ergs in visible light over a year or two were identified and defined by Fritz Zwicky and Walter Baade in 1934. A surprising amount of what they said then remains credible, including the division into Type I and Type II events (due, most astronomers now believe, to the release of nuclear and gravitational energy respectively), the association with rapidly expanding gas clouds (supernova remnants), and the association, at least sometimes, with collapsed objects at nuclear densities (neutron stars and pulsars), as outlined by R. Kirshner (University of Michigan) at a recent meeting*.

Type I events go off within relatively old stellar populations, in elliptical galaxies and between the arms of spirals. Their immediate progenitors are generally thought to be degenerate cores of C, O, Ne, Mg group elements, covered by a thin or thick envelope of helium. Such cores can arise either through mass transfer onto a white dwarf in a closed binary system or through stellar winds removing the hydrogen envelope from a 4–8 solar masses ($M_{\odot}$) star as it evolves. The explosions, reviewed by J.C. Wheeler (University of Texas), W.D. Arnett (University of Chicago), K. Nomoto (NASA-Goddard Space Flight Center), T.A. Weaver (Lawrence Livermore Laboratory), P. Sutherland (University of Texas) and S.R. Schurmann (University of Michigan), result from detonation or deflagration of the degenerate core, the fuel burning rapidly at high temperature. The dominant product is 0.2–1.4 $M_{\odot}$ of ^{56}Ni , which is radioactive and decays to ^{56}Co and then ^{56}Fe , with half lives of about 6 and 60 days respectively.

Gamma rays from these decays, greatly downgraded by scattering through the helium envelope, are then responsible both for the initial light outburst and for its exponential decay, also with a half life of about 60 days. The spectra of such events, reviewed by T. Axelrod (Lawrence Livermore Laboratory) and D.R. Branch

(University of Oklahoma), show no evidence of hydrogen, but essentially normal ratios of N, Mg, Na, S, Ca and so on, presumably embedded in helium, from the unburned envelope. After about 70 days, lines attributable to Co and Fe appear, the Co gradually giving way to the Fe on a time scale comparable with its half life.

This coherent picture is somewhat disturbed by the implications of X-ray observations of several galactic remnants. According to discussions by F. Winkler (Middlebury College), J.M. Shull (University of Colorado) and F. Seward (Center for Astrophysics), the remnants Pup A, Tycho and Cas A show excesses of intermediate weight elements (O, Ne, Si, and so on) but not of Fe; and the Tycho and Cas A remnants have masses near $15 M_{\odot}$. None of these, admittedly, is known with certainty to have been a Type I, but Tycho especially is often advertized as such.

Type II events occur only within very young stellar populations, such as in spiral arms. They apparently terminate the lives of massive stars (8–100 $M_{\odot}$) with extended ($\sim 10^{15}$ cm) hydrogen envelopes. The basic instability arises when nuclear burning produces an inert iron core in excess of the Chandrasekhar mass. Electron capture sets in, 'deleptonizing' the core and removing the degenerate electron pressure support that has been holding it up. Quite a modest amount of this is enough to send the central 0.6 $M_{\odot}$ or so of the core hurtling inwards. Neutrinos are trapped within the core and their degeneracy temporarily halts electron capture when the number of neutrons is roughly 1.4 times the number of protons (most of them still in rather exotic nuclei). As the inner core reaches nuclear densities, its equation of state suddenly stiffens, sending a shock wave back out into the remaining iron. This is called core bounce.

Just what happens next and how core bounce produces an observable supernova is both rather complex and rather uncertain. It was one of the special pleasures of the NATO meeting to hear several of the pioneers of nuclear astrophysics address these issues. H. Bethe (Cornell University) reviewed calculations

of the propagation of the shock wave and formation of a compact object. W.A. Fowler (Caltech) discussed electron capture cross-sections and shell-blocking, and their implications for deleptonization, and S.A. Colgate (Los Alamos) addressed the problem of fall-back of the outer stellar layers on to the neutron core. Other aspects of the problem were reviewed by Weaver, G. Brown (NORDITA), A. Yahil (State University of New York, Stony Brook), W. Hillebrandt (University of Munich), J. H. Applegate (Caltech) and S.E. Woosley (Lick Observatory). To make a proper Type II supernova light curve, spectrum and expanding remnant, the shock must get out through many solar masses of material, spread over $\sim 10^{15}$ cm, still carrying at least 10^{51} ergs. The hardest barrier to overcome is the remaining iron in the outer core, which extracts energy from the shock via photodissociation to helium and free neutrons. Rotation, magnetic fields and nuclear burning in the outer parts of the star (first discussed in 1960 by Fowler and Sir Fred Hoyle, who was also at the meeting) can all help, but whether or not the shock breaks out is a near thing, and must depend critically on details of the core structure. The answers may come from a new generation of two- (and three-) dimensional hydrodynamic codes now being developed in several places. Type II light curves are well modelled by the shock moving through an extended stellar envelope; and the spectra show the normal, hydrogen-rich composition of the envelope blown out by the shock at 10–20,000 km s⁻¹. The core presumably ends up as a pulsar, unless fall-back drives it over the stable mass limit for neutron stars into a black hole (perhaps most likely for the most massive progenitors).

Early in the history of the Universe, a third sort of stellar explosion may have occurred. Because the progenitors could have been very massive objects (100–500 $M_{\odot}$) essentially devoid of heavy elements, we might call them Type III supernovae. In

Virginia Trimble is a member of the Astronomy Program, University of Maryland, and the Department of Physics, University of California, Irvine.

*The NATO Advanced Study Institute on Supernovae was held 29 June–10 July 1981 at the Institute of Astronomy, Cambridge.

such objects, as discussed by Woosley, R. Bong (University of California, Berkeley), B. Carr (University of Cambridge) and W. Ober (University of Munich), the dominant instability results from electron-positron pair production in the hot, relatively low-density, core. Material expelled from such objects early in galactic evolution could plausibly raise the heavy element abundance of the remaining gas to 1–3 per cent of its present value, mostly in the form of intermediate-weight elements (O–Si) rather than Fe.

An urgent need in the field of supernova research is detailed observations of objects of both types just before and after maximum luminosity, in all spectral ranges. (Young extragalactic supernovae have now been seen as radio, X-ray and ultraviolet sources, as well as optical ones, according to K. Weiler (NSF), C. Fransson (NORDITA), N. Panagia (University of Bologna) and C. Canizares (MIT).) Early identification requires deliberate, organized supernova searches. R. Muller (University of California, Berkeley) and Colgate described the status of their respective automated search projects. Given adequate funding, both could be in operation within about 2 years, finding supernovae out to the Virgo cluster at a probable rate of several dozen a year. Most of these would be caught several days and several magnitudes before maximum luminosity. Such rapid discovery is particularly important if the Gamma Ray Observatory (with a projected life of only 2 years) is to probe nucleosynthesis by supernovae. E. Kibblewhite and M. Cawson (University of Cambridge) reported a slightly different kind of automated search, now in progress, that can find larger numbers of more distant events on photographic plates, also within a few days of maximum luminosity. Having a large, systematically acquired sample of supernovae will facilitate their use to determine the cosmological parameters H_0 and q_0 , via model atmosphere, standard candle, or Baade–Wesselink methods, outlined by R.V. Wagoner (Stanford University). This latter search should also help enormously to pin down the rates of the two types of event in different kinds of galaxy and, therefore, the nature of the progenitors.

Within our own Galaxy, rates of supernovae and their products now seem to be largely agreed upon, if not understood. G. Tammann (University of Basel) reported a total rate of 1 every 20 years (1:11 in a ratio 5:4) based on both historically observed events and extrapolations from the extragalactic rate. A.G. Lyne (Jodrell Bank Observatory) concluded that one pulsar is born every 28 years (probably a lower limit and subject to 50 per cent error) and D. Helfand (Columbia University) noted that the supernova remnant birthrate is only about 1 every 80 years (but is considerably higher, per unit mass, in the Large Magellanic Cloud). In addition, no

pulsar or neutron star is associated with many of the historical supernova, if neutron stars cool along the lines outlined by S. Tsuruta (University of Montana), though some intermediate-age remnants (RCW 103, CTB 109, G 21.5, G 74.9, 3C58, Vela and 4C 21.53W) do contain compact, probably non-thermal, X-ray sources, perhaps attributable to relativistic electron acceleration by a pulsar, according to Helfand and R.H. Becker (Columbia University).

The straightforward interpretation of these numbers and observations is that visible supernovae, remnants and pulsars can all be produced independently of the others, though two of the three must occur together frequently and all three occasionally, as in the Crab Nebula, which inspired Baade and Zwicky to some of their original suggestions. The assorted rates are reasonably consistent with the birthrates of the probable progenitors and with what is needed to build up the observed heavy element abundance of the Galaxy over its lifetime.

Each of the expanding supernova remnants has $\sim 10^{51}$ ergs ($\sim 10^{42}$ ergs s^{-1}) to dispose of. A few per cent of this, as discussed by R. Blandford (Caltech), goes into cosmic rays and other non-thermal sinks, but most of it eventually leaves the Galaxy as electromagnetic radiation. It does so by a rather devious route. R. Chevalier

(University of Virginia) and C. McKee (University of California, Berkeley) reviewed supernova remnant models in a way that made clear that the expanding remnants and the lumpy surrounding interstellar medium must be treated in a self-consistent way, including transfer of mass, energy and momentum between them. In particular, large, old remnants probably pervade a large fraction of the galactic disk, and should, therefore, be treated as a third, high-temperature, low-density component of the interstellar medium. This may sometimes erupt from the disk as a 'galactic fountain', according to F. Kahn (University of Manchester) and S. Ikeuchi (Hokkaido University).

Still further removed from the initial event, supernova explosions have been blamed collectively for triggering galaxy formation, as explicated by J.P. Ostriker (Princeton University), and the formation of the Solar System. D.D. Clayton (Rice University) presented an impressive body of evidence, drawn largely from chemical and isotopic abundances in meteorites, in opposition to the latter view. He believes that the cloud that collapsed to form our Sun did so as part of the normal, orderly evolution of the interstellar medium, which preserves in dust grains atoms synthesized in many supernovae over many billions of years. We are, indeed, made of starstuff, but very ordinary sorts of starstuff. □

Cholinergic neurones and memory

from Steven Zeisel, Daniel Reinstein, Suzanne Corkin, John Growdon and Richard Wurtman

A selective deficit in certain cholinergic neurones, possibly those in the septo-hippocampal tract, has been implicated in the aetiology of memory disorders associated with aging and Alzheimer's disease (presenile dementia). Agents that increase the release of acetylcholine from these neurones might therefore help to ameliorate the memory loss. This possibility provided a central theme for a recent workshop* on the pharmacology of memory disorders associated with aging. Of particular importance is the ability to determine in each patient whether his disease is, in fact, related to a central cholinergic deficiency, and whether acetylcholine is the sole or major

transmitter affected.

Data from cortical brain biopsies indicated that senile plaques and neurofibrillary tangles, both characteristic of Alzheimer's disease, were found in only 60 per cent of patients diagnosed using clinical criteria. These same 60 per cent also tended to show major decreases in brain choline acetyltransferase (CAT) activity, an index of the number of cholinergic terminals. Studies on autopsy specimens indicated that the density of central muscarinic cholinergic receptors apparently is unaffected in Alzheimer's disease, suggesting that patients might respond to compounds that increase cholinergic neurotransmission. Decreased levels of monoamine metabolites and somatostatin, a peptide neurotransmitter, in brain have also been reported; however, the levels of monoamine metabolites in cerebrospinal

*The International Study Group (ISG) on the 'Pharmacology of Memory Disorders Associated with Aging' held its second meeting in Zürich, Switzerland, on 3–5 April 1981. The 175 participants included basic neuroscientists, neurologists, psychiatrists, neuropsychologists, administrators of granting agencies and research scientists at food and drug companies. The proceedings of the meeting will be published by Raven Press, New York, in *Alzheimer's Disease: A Report of Progress in Research* (eds S. Corkin, K.L. Davis, J.H. Growdon, Earl Usdin and R.J. Wurtman).

Steven Zeisel, Daniel Reinstein, Suzanne Corkin, John Growdon and Richard Wurtman are in the Department of Nutrition and Food Science, Massachusetts Institute of Technology.

fluid are not sufficiently altered to aid in diagnosis. Methods, including computerized axial tomography and positron emission scanning, for identifying patients with Alzheimer's disease without resorting to cortical biopsies were also discussed.

The loss of cholinergic neurones associated with Alzheimer's disease and the abundant pharmacological evidence that anticholinergic compounds impair memory functions even in normal people suggested that enhancement of central cholinergic neurotransmission might aid dementia patients. This hypothesis has been difficult to test, because no Federal Drug Agency approved, orally active, non-toxic cholinergic drugs exist. Another strategy for increasing central cholinergic tone involves using nutrients, such as choline or the choline-containing dietary phospholipid lecithin, to accelerate the synthesis of acetylcholine in, and its release from, brain neurones. The response of a cholinergic neurone to more precursor depends on its firing frequency: when firing rapidly, neurones are highly responsive, but when firing slowly, they respond little, if at all, to more choline. This frequency dependence allows exogenous choline (or lecithin) to act selectively, without causing generalized manifestations of cholinergic hyperactivity. It might be anticipated that in diseases like Alzheimer's, where most of the neurones in particular cholinergic tracts have disappeared, those surviving would be hyperactive and thus choline responsive.

The difficulties in assessing memory function in elderly subjects were discussed in detail. A concurrent *ad hoc* workshop on the classification of the severity of dementia concluded that no simple, global rating scale could be proposed yet. It did, however, suggest guidelines for developing such a scale. Experimental studies discussed during the first ISG meeting (1-2 December 1979) had suggested that agents that enhance central cholinergic tone might be beneficial in Alzheimer's disease. Since then, numerous investigators have initiated well controlled clinical trials using choline, lecithin, physostigmine and other agents thought to increase central cholinergic tone. In the approximately 100 patients with Alzheimer's disease discussed by various groups at the meeting, administration of choline or lecithin alone had no reproducible beneficial effect. Preliminary data from small studies, using the direct muscarinic agonist physostigmine or combinations of choline or lecithin with oral physostigmine or piracetam, appear more promising than the results with lecithin alone. (The possibility that choline or lecithin might slow the deterioration of mental status associated with Alzheimer's disease was suggested by studies showing that long-term choline administration to aged rodents suppressed the age-related loss of cortical dendritic spines and preserved retention of passive-avoidance learning.) L-dopa, the peptides lysine vaso-

pressin, 1-desamino-8-D-arginine vasopressin, 9-desglycinamide-8-L-arginine vasopressin, or ACTH (4-9) had modest or no therapeutic effects in patients with Alzheimer's disease.

The design of clinical studies to test the effects of any agent on cognitive function in patients with Alzheimer's disease is hampered by the fact that groups of clinically similar patients are bound to be neuropathologically and neurochemically heterogeneous. For example, some of these patients will lack senile plaques and neurofibrillary tangles, and others will have normal CAT activity. There is, however, no way to identify more homogeneous subgroups in clinical practice. If, for example, only 20 among 100 clinically similar patients have a selective loss in cholinergic neurones, the others either lacking such pathology or also having deficits in somatostatin, the monoamines, or other neurotransmitters, then only those 20 would be expected to respond to cholinergic stimulation. However, as there is no way of detecting those 20, a treatment

improving 3 out of 4 of them would, in fact, affect only 15 per cent of the treated population and might falsely be labelled ineffective. One way to get around this problem is to identify 'responders' and see whether they respond repeatedly. In addition, non-invasive methods for assessing the release of brain acetylcholine and other neurotransmitters need to be developed if we are to design rational, neurochemically based treatments for Alzheimer's disease. Such techniques might also be of use in understanding other neurological syndromes that arise from multiple and distinct pathological processes.

The meeting yielded no treatment for Alzheimer's disease, but it did confirm the progress made in measuring memory functions in afflicted people, in characterizing their neurochemical deficits and in developing a rational basis for an approach to treating at least one of these deficits, that in cholinergic brain neurones. The meeting also generated a consensus as to new lines of research that need to be explored. □

How different are human races?

from J.S. Jones

It is generally accepted that the human population can be divided into groups which differ from each other in obvious characteristics such as skin colour. The classical (and most popular) views of race are *typological*; they imply that each race is a homogeneous group of individuals, and that the members of such a group are biologically similar to each other, and different from the members of other racial groups. Each race is assumed to represent a genetically differentiated human type.

There are serious problems in defining races in this way, as it is now obvious that each human population, rather than consisting of a uniform group of individuals, contains within itself a large amount of genetic variation for many of those characters originally used to define racial 'types'. Faced with this variation, supporters of typological views have been forced into increasingly unlikely models involving the mixture by migration of originally pure populations. Physical anthropology has found it difficult to come to terms with genetic polymorphism. This very diversity can, however, be used to provide information for an estimate of the differences among human populations.

Rather than defining race on the basis of a few conspicuous variants such as skin colour (which may be based on genetic differences of as few as four loci), it is now possible to

use information on the distribution of genetic polymorphism at scores of loci for blood groups, histocompatibility antigens, enzymes and the structure of DNA revealed by restriction enzymes. Some of the discontinuities among populations for these newly discovered polymorphisms are as great as those which exist for skin colour; the Duffy blood group system, for example, has an allele in 100 per cent of the populations of the Pacific Islands which is absent from South American Indians².

Even a superficial examination of the geographical patterns of allele frequency in man shows that the typological view of each race as a genetically well differentiated group is not correct. The geographical trends of gene frequency for a sample of human polymorphisms hardly ever parallel those for skin colour or body form. New statistical techniques which combine all the available information on biochemical polymorphism make it possible to quantify the extent of genetic differentiation within and among human populations and give a new insight into the nature of race.

Latter³ has developed an index which uses data on 18 polymorphic gene loci from 180 different human populations from each of the six major racial groups (classified as European, African, Indian, East Asian, New World and Oceanian) to give a measure of the proportion of genes which two randomly chosen individuals have in common. Comparison of the value

J.S. Jones is in the Department of Genetics and Biometry, University College London.

The Burgess Shale: a unique Cambrian fauna

from R.A. Fortey

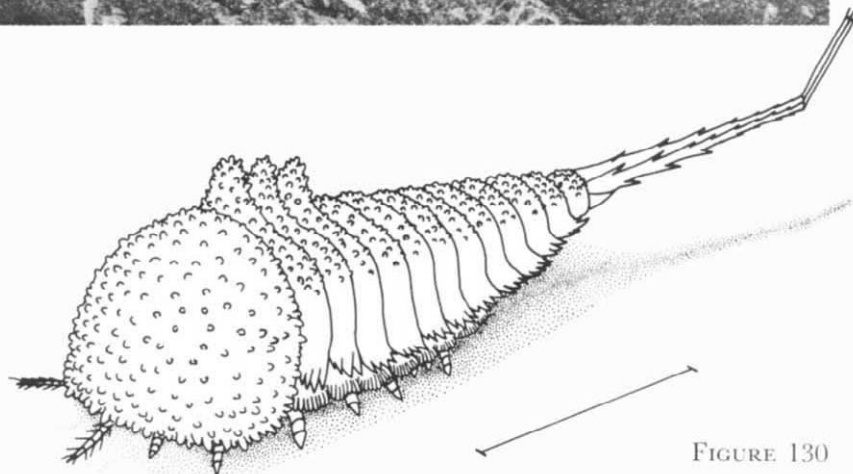


FIGURE 130

Habelia optata. Above, fossil from the Walcott quarry. $\times 4.6$. Below, reconstruction in oblique anterior view, suggesting the animal walking on the sea bottom.

Molaria spinifera. Fossil from the Walcott quarry. $\times 5$.

THE fossil record is usually very selective, preserving only the remains of well skeletonized animals. Soft tissues decay and leave no trace, and many groups of organisms lacking tough, preservable shells or skeletons have virtually no fossil record as a consequence. There are extremely rare cases where the rocks have preserved impressions of delicate limbs or internal organs, as well as of different kinds of worms and other organisms otherwise unknown as fossils. The importance of such occurrences is enormous in fleshing out the full complexity of organic evolution.

Probably the most significant of these 'geological miracles' is the Burgess Shale in British Columbia, of middle Cambrian age (about 550 million years BP). The locality was originally described by the great American palaeontologist C.D. Walcott in the earlier part of this century.

Walcott named a whole array of amazingly well preserved fossils: 'worms' and sponges, and an astonishing variety of arthropods, as well as a number of enigmatic fossils which could not be assigned readily to any known group of animals. This fauna is now under a thorough re-examination (with much new material from British Columbia) by a team of scientists led by Professor H.B. Whittington of Cambridge University. The illustrations here are from the latest, and one of the last, parts of this reconsideration.

What is most surprising about the new results is that the puzzles have, if anything, multiplied. Very few of the arthropods can be simply related to groups still living. More often they show curious combinations of features — many of them primitive, but some specialized in odd ways — that make their

classification a delicate business, to say the least. Some even have body segmentation resembling one group of arthropods, but limbs more like another! It is certain, however, that the arthropods must have had a complex history before the middle Cambrian. Some of the 'worms' are members of living groups: in one case their Cambrian diversity was far greater than at the present day. But there remain a number of bizarre, soft-bodied animals which defy ready classification with any of the major groups of living animals. There has even been the suggestion that these are the lone records of phyla of animals which have not survived — unless they are still lurking undiscovered somewhere in our oceans.

R.A. Fortey is in the Department of Palaeontology, British Museum (Natural History).

of this index obtained when the two individuals come from the same 'race' with that obtained when the two are members of different 'races' gives a clear indication of the degree to which the human population is divided into genetically different groups. His analysis shows that by far the largest component of the total genetic diversity of mankind — about 84 per cent of all genetic variation — results from the genetic

differences which exist between individuals belonging to the same tribe or nationality. About six per cent arises from differences between tribes or nationalities (such as those which are found, for example, between the populations of France and Spain, or between different tribes in the east and west of Africa). Only about ten per cent of the total biological diversity of mankind arises from genetic divergence

between 'racial' groups. In other words the genetic differences between the classically described races of man are on the average only slightly greater than those which exist between nations within a racial group, and the genetic differences between individual human beings within a local population are far larger than either of these. Mankind as a whole, far from being divided into a number of discrete entities which could be

defined as racial 'types', is in fact rather a homogeneous species in terms of geographical variation. Support for this conclusion comes from other statistical tests of the distribution of blood group and enzyme polymorphisms^{4,5}, and from some very preliminary evidence on sequence polymorphism in human mitochondrial DNA; the two most similar of 21 individuals characterized for this polymorphism were of African and European origin respectively⁶.

The genetic evidence also favours a recent origin of human racial groups. Populations which are nearest the global mean in their frequency of enzyme variants and blood groups cluster in south-central and eastern Asia⁷; these may represent the ancestral populations of *Homo sapiens*, from which waves of migration (coinciding with the origin of agriculture) colonized the rest of the world. Assuming, with some optimism, that enzyme and blood group variants are on the average not subject to strong natural selection, and that they originate at a fairly uniform rate among all mammals, the probable time of divergence among the resulting populations of Africa, Europe and the Orient is only about 60 — 120,000 years before the present day⁸, a period very short in terms of the evolution of the primates⁹. Estimates of the rate of accumulation of mutations in mitochondrial DNA in man compared with other primates give a similar date of divergence among the human populations of the world^{6,10}.

Although the races of classical anthropology and popular myth do not represent genetically differentiated entities, it is of course possible to use the geographical variations in gene frequency which do exist as the basis of a taxonomy of mankind. Several trees showing the evolutionary relationships of various population groups have been constructed. Until recently there was considerable disagreement between those based on skeletal characters and those based on molecular polymorphisms. It now appears that this arises largely from the strong correlations which exist between skull form and climate. These lead to a convergent evolution of populations

living in similar climates great enough to obscure the patterns due to common ancestry. Once these climatic effects have been removed by statistical means, both skeletal and biochemical measures of identify produce trees which show that Africans and Europeans are more closely related to each other than either is to Amerindians or to Australasians¹¹.

As well as using genetic data to investigate how much human groups have in common, it is possible to use such information to ask how differentiable the various groups might be. Such techniques of identification have long been used by forensic scientists who use man's extensive within-population diversity to assess the probability of a blood sample coming from a particular individual. As the probability of two samples from the same population having an identical genotype at a polymorphic locus is always less than one, it is inevitable that, as more loci are considered, it will become possible to identify single members of the population with great accuracy. In the same way, that relatively small proportion of the total human diversity which arises from genetic differences between populations can provide an estimate of the population from which a particular blood sample has been taken; using a multiplicative measure of

identity on as few as eight enzyme or blood group loci will permit the assignment of a single blood sample to its correct racial group with at least an 87 per cent probability of accuracy¹².

A confusion between these two measures of genetic identity — the proportion of genes which members of different races have in common, versus the probability of being able to identify to which race an individual belongs — has led to a recent controversy about the real extent of the genetic differences in the races of mankind. Mitton¹³ claims that, as by multiplying the probability of identity of individuals of different racial origin for one polymorphic locus after another it is possible to assign them with some precision to the racial group from which they come, then human races must each represent a rather major level of genetic divergence. However, the index of identity used by Mitton is a multiplicative one; it does not give information as to the extent to which races show overall genetic differences among themselves. Several authors^{14–16} have pointed out that Mitton's measure of the genetic identity of populations cannot be used to assess the average genetic differentiation of human races. The idea of racial 'type' — and, some would argue, of 'race' itself — is no longer a very useful one in human biology. □

Early results from Magsat

from R.A. Langel, R.H. Estes and M.A. Mayhew

NASA's Magsat satellite programme was designed to study the near-Earth magnetic field. Specifically, its objectives were to describe the main field (presumably) originating in the Earth's core and to compile and interpret the much smaller fields due to variations in the magnetization of the lithosphere. The 1981 spring meeting* of the American Geophysical Union, held on 27 May in Baltimore, Maryland, offered the first opportunity for discussion of the early results.

Magsat was launched on 30 October 1979 into a Sun-synchronous (twilight) orbit with apogee of 561 km and perigee of 352 km. Re-entry occurred on 11 June 1980. The instrument complement included a caesium vapour magnetometer to measure the field magnitude, a fluxgate magnetometer to measure the field components and an optical system to measure the orientation of the fluxgate magnetometer. The scalar magnetometer was used to calibrate the vector magnetometer in flight. The data are estimated to be

accurate to about 3 nT in magnitude and 6 nT in each component.

After a description by Langel (Goddard Space Flight Center) of the Magsat mission and data, participants discussed spherical harmonic models and fields due to ionospheric and magnetospheric currents. Estes (Business and Technological Systems) described a series of models incorporating data from observatories, repeat stations, marine surveys and the POGO satellite, as well as Magsat, in various combinations. He emphasized the value of a new technique which simultaneously solves for the anomaly field at each observatory and for the main-field coefficients. Using this technique, it is now practical to include third-time derivatives of the spherical harmonic coefficients without serious misbehaviour of those coefficients. Such models seem to predict better than those not using third derivatives and not solving for observatory anomalies for a period of 3–4 years; after that time interval the

R.A. Langel is a Magsat Project Scientist (Geophysics Branch) at the NASA-Goddard Space Flight Center and R.H. Estes and M.A. Mayhew are Magsat Investigators at the Business and Technological Systems, Inc., Maryland.

*Abstracts of all papers presented at the meeting may be found in *EOS, Trans. AGU* 62, 268 (1981). The Magsat studies are being undertaken by the Geophysics Branch at NASA's Goddard Space Flight Center, by the US Geological Survey and by investigators from the United States and elsewhere (see Langel *Johns Hopkins APL Technical Digest* 1, 214 (1980)).

1. Stern, C. *Human Hered.* 20, 165 (1970).
2. Mourant, A.E., Kopen, A.C. & Domaniewska-Sobczak, M. *The Distribution of the Human Blood Groups and other Polymorphisms* (Oxford U.P., 2nd edn, 1976).
3. Lattier, B.D.H. *Am. Nat.* 116, 220 (1980).
4. Lewontin, R.C. *Evol. Biol.* 6, 381 (1972).
5. Nei, M. & Roychoudhury, A.K. *Am. J. hum. Genet.* 26, 421 (1974).
6. Brown, W.M. *Proc. natn. Acad. Sci. U.S.A.* 77, 3605 (1980).
7. Piazza, A., Menozzi, & P., Cavalli-Sforza, L.L. *Proc. natn. Acad. Sci. U.S.A.* 78, 2638 (1981).
8. Nei, M. *Jap. J. hum. Genet.* 23, 341 (1978).
9. Bruce, E.J. & Ayala, F.J. *Evolution* 33, 1040 (1979).
10. Ferris, S.D., Wilson, A.C. & Brown, W.M. *Proc. natn. Acad. Sci. U.S.A.* 78, 2432 (1981).
11. Guglielmino-Matessi, C.R., Gluckman, P. & Cavalli-Sforza, L.L. *Am. J. Phys. Anthropol.* 50, 549 (1979).
12. Smouse, P.E. & Spielman, R.S. in *Human Genetics* (eds S. Armadures & Lisker, R. Excerpta Medica, Amsterdam, 1977).
13. Mitton, J.B. *Am. Nat.* 111, 203 (1977).
14. Chakraborty, R., et al. *Am. Nat.* 112, 1134 (1978).
15. Neel, J.V. *Am. Nat.* 117, 83 (1981).
16. Nei, M. *Am. Nat.* 117, 88 (1981).

higher derivatives begin to dominate and the models are untrustworthy. A model for 1950–1980 describes extremely well the secular variation for that time period. Furthermore, pre-Magsat models of this nature were the best predictors of the fields measured by Magsat. A preliminary model based on only 15 magnetically quiet days of Magsat data was also presented.

New areas in field-modelling methods are being investigated by Gibbs (Business and Technological Systems) and by Shure (University of California, San Diego). The former presented a new approach to geomagnetic core field modelling using optimal recursive filtering techniques. The filter has as its state variables the spherical harmonic coefficients of the scalar potential and its first derivatives obtained from conventional least-squares analysis over a deterministic interval of approximately 5 years. Using statistical information about the temporal variation in the coefficient derivatives to weight the data (the coefficients of the deterministic 5-year solutions), the technique was applied using five models covering the interval 1950–1976 to predict the magnetic field in 1980. Prediction of the Magsat data using the filter technique was better than with other pre-Magsat main-field models.

Shure has developed a method of applying constraints designed to produce 'smoothed' models at the Earth's core, that is, models of the core field where the effect of short-wavelength contamination from crustal and external fields is minimized. Using the smoothness assumption of minimizing the total magnetic field energy outside the core, the continuous inverse problem was formulated using spherical harmonics as a basis of spline functions on the sphere. The matrix elements are constructed in closed form, and the fields at the core–mantle boundary are convergent. The method requires the inversion of a matrix of dimension equal to the data set. A numerical simulation was described based on a geographical distribution of 21 observatories on the Earth's surface and using data generated from an IGRF model. The smooth field was examined at the Earth's surface and at the core–mantle interface. Because of the size of the matrix to be inverted, the method is computationally somewhat impractical. Efforts are being made to adapt it for use with large data sets.

Several speakers (Cain, US Geological Survey; Zanetti, Johns Hopkins University; Olson, McDonnell Douglas Astronautics Co.; Hermance, Brown University; and Sugiura, Goddard Space Flight Center) emphasized that the effects of ionospheric and magnetospheric currents must be taken into account when modelling the Earth's main field and when studying crustal anomalies. Such effects include unbalanced field-aligned currents which result in significant fields far from the current system, inaccuracies in magnetospheric models near the Earth,

fields induced from time-varying sources and the ever-present quiet daily current system, Sq.

Hughes (Herzberg Institute of Astrophysics), on the other hand, showed that, when studying quiet time current systems, it is important to take into account the existence of fields from crustal sources. When the anomalous fields are accounted for, the quiet Magsat data from November 1979 can be shown to be in qualitative agreement with a model consisting of field-aligned currents and their closing currents within the ionosphere.

Using quiet Magsat data from November 1979, Wilson (Herzberg Institute) showed that the magnitude of the field-aligned currents in the sunlit hemisphere is greater than in the dark hemisphere, and that there is considerable structure in field variation over the summer polar cap which is absent in the winter hemisphere.

Taylor (Goddard Space Flight Center) discussed a very important study showing that inaccuracies in the gravity field model used to derive the spacecraft ephemerides are unlikely to produce satellite apparent motions which would result in 'false' magnetic anomalies, because the wavelengths of effects with amplitude sufficient to be significant to the satellite magnetic data are of the order of the circumference of the Earth and can be filtered out of the data. Simulations pertinent to Magsat showed that the effects of orbit error on that data are insignificant in all cases.

Further papers were concerned with identifying and interpreting fields from lithospheric sources. Langel briefly reviewed the history of satellite-derived magnetic anomaly maps, beginning with that of Regan *et al.* (*J. geophys. Res.* **80**; 794), published in 1975, and culminating in the latest map derived from Magsat data. The procedure for separation of the crustal fields from the Earth's main field and from external fields is the following (Regan *et al. op. cit.*; Mayhew *J. Geophys.* **45**; 119, 1979; Langel *et al. Can. J. Earth Sci.* **17**; 876, 1980): (1) select data from periods of relative magnetic quiet, (2) remove the core field estimated by a thirteenth degree and order field model, and (3) filter the data appropriately along the orbit track to remove the remaining long-wavelength effects of core and external fields. This filtering technique introduces a zero-level ambiguity and artificial east–west striping in the resulting maps. Future research will investigate alternative techniques which minimize these shortcomings. The main features of the Magsat anomaly map are the same as those found in the higher-altitude POGO data, but many new details are present.

Frey (Goddard Space Flight Center) pointed out many correlations between known large-scale tectonic features and the satellite-derived anomaly maps, including positive anomalies over many major shields, cratons, blocks and stable basins, a



100 years ago

The Yorkshire Naturalists' Union will have a Fungus Foray on Friday and Saturday, September 30 and October 1, at which they will gladly welcome any mycologists who may be disposed to assist them. The Friday's programme is to consist of an excursion in the neighbourhood of Harrogate. On the Saturday is to be a "show," at which will be exhibited fungi, and any objects illustrative of the subject which may be sent. The dinner is to be on the evening of Saturday. Arrangements are being made to search localities in all parts of Yorkshire for specimens to exhibit.

From *Nature* **24**, 502, 22 September, 1881.

strong negative anomaly following the topography in the Himalayan–Tibet region and many anomalies bounded by what are thought to be plate-tectonic sutures and failed rifts. The high degree of correlation between known geology and the satellite magnetic anomaly map was reinforced by Hastings (Technicolor Graphic Services, Inc., South Dakota) regarding the African continent and by Sailor (Analytical Sciences Corp., Massachusetts) in a study of the south-east Indian Ocean. Sailor showed that Magsat anomalies are correlated with tectonic features (Ninety East ridge, Diamantina Fracture zone, Broken ridge) and from a spectral analysis concluded that the early Magsat data have a 'noise floor' of 1–2 nT at a wavelength of ~ 360 km.

Wasilewski (Goddard Space Flight Center) examined data on xenoliths, high-grade metamorphic rocks and a suite of rocks across the Ivrea zone. These are presumed representative of lower crustal and upper mantle petrology. He concluded that the source of long-wavelength magnetic anomalies over continental areas is the lower crust, that the mantle is non-magnetic and that the Curie point is in general about 560°C, except in certain active zones, such as the Rio Grande Rift, where it may be considerably lower.

Hinze (Purdue University) and Phillips (US Geological Survey) compared aeromagnetic data with the satellite anomaly maps for the United States. The aeromagnetic data were first extrapolated upwards to satellite altitude. Both speakers, but particularly Hinze, found remarkable agreement in many features. For some unknown reason, Phillips found some anomalous areas in the midwest showing poor correspondence with both the satellite data and the results of Hinze. von Frese presented a geological model of the Mississippi Embayment which reproduces the major features of both the long-wavelength gravity and magnetic anomalies in that area and is also in agree-

ment with the known seismic refraction and heat flow data. His group believes that the region is the site of a late Precambrian rift which was reactivated in the Mesozoic to form an aulacogen and its model involves thinning of the upper crust and an increase in density in the lower crust. The magnetic low is accounted for by either an isotherm upwarp or, less likely, a lithological variation in the crust. Mayhew's

(Business and Technological Systems, Inc.) magnetization model for the United States, derived from Magsat scalar data, had much better resolution than that using POGO data.

Although preliminary, the results discussed at the meeting indicate that separation of the measured field into its core, crustal and external 'components' is being achieved. The newest discipline, inter-

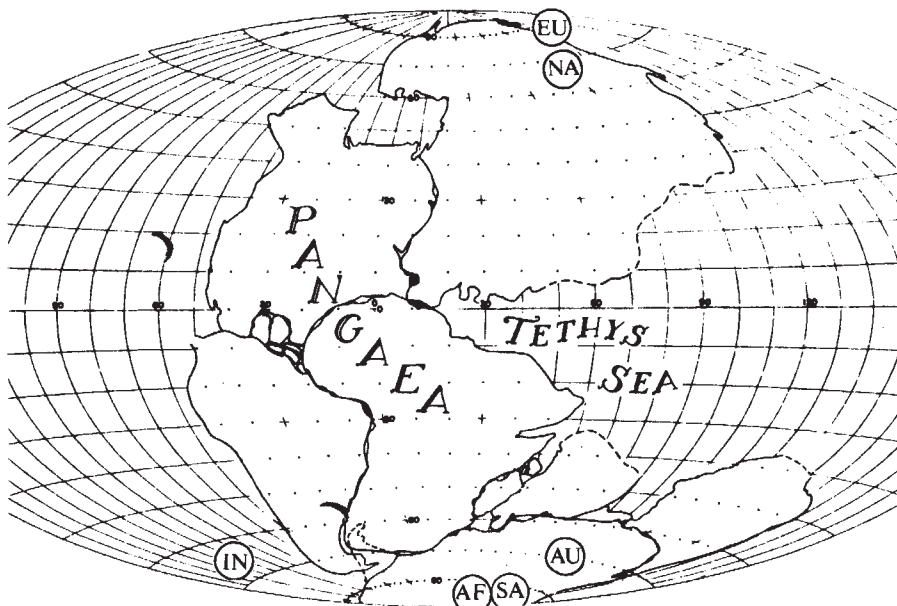
pretation of long-wavelength crustal fields in terms of a geological/geophysical model, promises to contribute significantly to our understanding of the Earth's crust. In the older disciplines of main-field modelling and studies of external fields, there are significant new developments in both analytical techniques and in our understanding of the physics of the field sources. □

The East Asian jigsaw puzzle — Pangaea at risk?

from Neville Haile

UNTIL fairly recently, reconstructions of the world palaeogeography followed Wegener in showing Eurasia, excluding the Indian subcontinent, as a single block, with the Malay Peninsula and part or all of the Indonesian Archipelago depending from it and looking rather vulnerable^{1,2}. Other world palaeogeographical maps simply omit South-east Asia and most of China (see the figure)^{3,4}.

The suggestion that South-east Asia, or part of it, was initially separate from mainland Asia, to which it later became welded, dates back at least to Wing Easton⁵ in 1921. More recent variations show South-east Asia (with or without parts of China) in Palaeozoic or early Mesozoic times attached to various parts of Gondwana, or to Laurentia in the same configuration as at present, or detached from both^{6,7}. Argand's⁸ idea that the rest of mainland East Asia south of the Siberian craton may be an aggregate of a number of blocks, which have had a separate history and become welded together after colliding, has been recently revived^{9,10} on the basis of new and more detailed maps produced by Chinese geologists, showing belts of calc-alkaline rocks and ophiolites interpreted as fossil subduction zones¹¹. A recent Sino-American review¹² asserts: "In a nutshell, China is the product of the processes which led to gradual accretion onto the Siberian craton of a number of island arc systems and the capturing of relatively small continents. This process continued with the collision of the Indian subcontinent with Asia". Palaeomagnetic methods are ideally suited to test these hypotheses, delineate the pieces of the jigsaw puzzle and trace their movements through geological time, but hitherto palaeomagnetic evidence from China has been sparse and inconclusive. Moreover, this lack of information from China makes it difficult to interpret and integrate the data available from the peripheral areas of Siberia, Korea, Japan and the Indian subcontinent, and South-east Asia¹³.



Reconstruction of the continents into the universal landmass of Pangaea as of the end of the Permian, 225 Myr ago — the 'classical' concept of Wegener in a version by Dietz and Holden³. Where were East and South-east Asia?

The paper by McElhinny *et al.* (this issue of *Nature*, page 212) is significant in showing that rocks from the same system (Permian) on two of the postulated tectonic blocks of China show different palaeomagnetic vectors, supporting the idea that in the Permian they were separated from each other and from other parts of Asia for which there is published data. The authors' geographical reconstruction for the Permian shows four continental blocks drifting independently in a proto-Pacific Ocean, waiting to collide with and be welded to the mainland (India and Qinhai-Tibet being off-stage, presumably to the south, forming part of Gondwana). If this is correct, then Wegener's percipient idea, now orthodox theory, that all the continents were joined into a single supercontinent of Pangaea in the Permian, is an over-simplification.

The authors' model is only one of several that could satisfy the data. Ideally the next step should be an attempt to establish palaeomagnetic vectors from all geological systems over wide areas of East and South-east Asia in order to define

apparent polar wander curves for each of the postulated blocks, and to integrate the results with evidence from structural geology, palaeoclimatology, fossils and (for the late Mesozoic and Cenozoic) ocean-floor spreading history. To achieve this would seem to call for a major cooperative effort — an international decade of palaeogeography? □

1. Wilson, J.T. *Sci. Am.* **208**(4), 86 (1963).
2. Smith, A.G. & Briden, J.C. *Mesozoic and Cenozoic Palaeogeographical Maps* (Cambridge University Press, 1977).
3. Dietz, R.S. & Holden, J.C. *J. geophys. Res.* **75**, 4939 (1970).
4. Irving, E. *Nature* **270**, 304 (1977).
5. Wing Easton, cited by du Toit, A.L. in *Our Wandering Continents — an Hypothesis of Continental Drifting*, 126 (Oliver and Boyd, Edinburgh, 1937).
6. Scotese, C.R., Kambach, R.K., Barton, C., van der Voo, R. & Ziegler, A.M. *J. Geol.* **87**, 127 (1979).
7. Haile, N.S. *Earth planet. Sci. Lett.* **48**, 233 (1980).
8. Argand, E. *C. r. 13th int. geol. Congr.* 1922, Brussels, **1**, 171 (1942).
9. Burret, C.F. *Earth planet. Sci. Lett.* **21** 181 (1974).
10. Terman, M.J. *Abstr. Bull. Am. Ass. petrol. Geol.* **62**, 12, 2513 (1978).
11. Huang, Chi-Ching (T.K.) *Eclog. geol. Helv.* **71**, 611 (1978).
12. Bally, A.W. *et al.* *U.S. geol. Survey Open-file Rep.* 80-501, Reston, Virginia (1980).
13. Haile, N.S. in *Paleoreconstruction of the Continents*, 129 (American Geophysical Union Geodynamics Series 2, 1981).

Neville Haile is at Robertson Research International, Llanrhos, UK.

ARTICLES

Glacial and interglacial wind regimes over the eastern subtropical Atlantic and North-West Africa

M. Sarnthein*, G. Tetzlaff†, B. Koopmann*, K. Wolter† & U. Pflaumann*

* Geologisch-Paläontologisches Institut der Universität Kiel, D-2300, FRG

† Institut für Meteorologie der Universität Hannover, D-3000, FRG

Analysis of aeolo-marine dust deposits in the subtropical eastern Atlantic enables the strength of the major wind patterns during the late Quaternary to be evaluated and gives an insight into the climate of North Africa.

THE origin of the late Quaternary climatic changes in the subtropical arid belt of North Africa is still unknown. Of the explanations of the relationship between subtropical and global climatic conditions during glacial and interglacial times one¹⁻⁴ suggests that there is a fairly constant latitudinal extension of the desert belt during both the warm and cold stages, accompanied by distinct latitudinal shifts which are directly related to the expansion and shrinkage of the north polar ice cap. Another approach⁵⁻⁷ assumes that the circulation of the atmosphere over subtropical North Africa has a basically stable latitudinal position but that there are considerable variations in the intensity of circulation. Accordingly, the centre of the arid belt would have remained in the same position, but the belt would have expanded during the cold periods and contracted during the warm ones. New evidence is provided by our analysis of aeolo-marine dust deposits in the subtropical eastern Atlantic, which evaluates grain-size data from deep-sea samples using two approaches. The results suggest a stable location of the aeolian trajectories. They also enable the strengths of the major wind patterns that existed during the late Quaternary to be quantified, which has important implications in understanding the climate of North Africa.

Meteorological setting

The major dust outbreaks at the continental margin of West Africa occur between June and September and are associated with warm, dry plumes of Saharan air. They are located in the easterly midtropospheric jet stream which we term 'Harmattan'⁸ or the 'Saharan Air Layer'⁹⁻¹¹. This dust flow is linked with the 'Easterly Waves', which are a system of tropical disturbances connected to the intertropical convergence zone (ITCZ) and appearing with a 3-4 day cycle^{10,12,13} (Fig. 1a). The dust plumes are generated by gusty winds and strong convective mixing. These processes are caused both by surface heating and the north-flowing wind component in the northeastern sector of the Easterly Waves^{12,14}. Such conditions are restricted to the south Saharan and Sahel zones. Here, the widespread lateritic or ferruginous crusts (Fig. 1a)^{15,16} produce a characteristic dust which is composed of predominantly kaolinite in the clay fraction^{17,18} and abundant reddish-stained quartz¹⁸⁻²⁰.

Having crossed the ITCZ, the dust cloud flows in a westerly direction with the Harmattan air, and its maximum concentration is at a height of ~3,000 m. Beyond the shore line, the dust moves above the Trade Wind inversion (at 600-1,500 m) circulating round the upper-air high-pressure system which is found over the western Sahara and is shown in Fig. 1a. This leads to a sickle-shaped dust trajectory over the eastern Atlantic. The model proves to be correct for 75% of the data contained in a 15-yr meteorological survey¹³. The fine-grained components of the dust, which have not yet fallen out, finally move westwards to the Antilles¹¹. The zone of gravitational fallout in the south is

limited by the northward penetration of the monsoonal rains.

Additional dust pulses advance in the wake of the Easterly Waves¹¹. The shallow Trade Wind layer below the Harmattan flow has also been considered responsible for the Saharan dust supply^{21,22,49}; but this wind system contributes a heavy dust load only to a limited offshore range (Fig. 1b)^{19,23}. These sediments originate from the northwestern fringe of the Sahara; they are characterized by the prevailing chlorite, illite in the clay fraction and pale quartz^{17,18,20}. Furthermore, the Trade Winds tend to convey the fallout from the Harmattan dust plumes towards the south-west. The dust load supplied by the Trade Winds during the winter²⁴ is not well recorded by the sediments in the relevant parts of the equatorial Atlantic^{19,25} and obviously has minor importance.

Sediment distribution

The modern dust fallout over the subtropical area of the North Atlantic Ocean is documented by 84 sediment-surface samples, each of which represents a record of the past 500-2,000 yr. The terrigenous sediment fraction >6 µm was analysed (on a carbonate and opal free basis) for 15 0.33 phi-grade size fractions²⁶. The grain-size percentage values, which were weighted for their respective proportions of fine silt and clay <6 µm, were row-normalized and Q-mode, rotated principal components (factor) analysed²⁷. The resultant grain-size assemblages (factor numbers 1, 2, 3, and 4) which account for 99% of variance, are centred around 11, 24, 74 and 45 µm modal sizes respectively. Although the terrigenous fraction offshore West Africa has at least three major sources (the loads from the Harmattan and Trade Winds and the rivers), the Harmattan dust dispersal is the dominant control underlying the regional distribution of the grain-size factors 1, 2, and 4. Factor 3 (4.5% of variance) seems to reflect bottom-water transport (characteristic samples see Fig. 2).

The weighted silt grain-size assemblage of factor 2 (24-µm mode) is characteristic for sediments formed by loess-like proximal dust falls which occur at distances of <800-1,000 km from the continental source⁹ (Fig. 1b). The sickle-shaped maximum of factor 2 at 18°-26° N matches the similar maxima of the grain-size fraction >6 µm (ref. 19), of quartz²⁵ and of kaolinite^{17,18}. It is highly unlikely that bottom currents produced these features. The sickle pattern does not match any of the mud wave zones reported for this general area^{28,29}. It can be assumed, therefore, that the sickle shape corresponds to the prevailing trajectory of dust plumes in the Saharan Air Layer during the summer. In the factor 2 pattern the centre of the dust outbursts from the coast apparently corresponds to a narrow latitude band, 16°-21° N, which would mark the present average northernmost position of the ITCZ. Note that transportation in both the surface Trade Wind layer and the oceanic currents only slightly affects the distribution of the Harmattan-supplied dust

in the deep-sea sediment. The sickle-shaped depositional patterns derived from different grain sizes¹⁴ are largely congruent and do not depict a strong influence of the Trade Wind direction. Presumably, the ineffectiveness of ocean transport is largely due to a coarsening of the dust grain sizes close to the sea surface because of the formation of faecal pellets³⁰.

The pattern of factor 2 also reflects the depositional patterns of the complementary fine-sediment fractions (factor 1: 11- μ m mode). These are derived from two sources. First, the dust discharge contributes the predominantly fine fractions which remain after long-distance transport, both to the more remote areas of the Atlantic and to a sector 21°–24° offshore. Second, they are derived from fluvial sediments close to the continental margin (Fig. 1b). River supply is indicated, for example, by an excess of fine-grained material <6 μ m in the basically bimodal grain-size distributions^{18,26}, such as those occurring in small-scale, lobate areas off the coasts of Senegal, Guinea and south Morocco. Complementary evidence comes from abundant plant fibres and well preserved faecal pellets in the coarse fraction of the sediment.

However, the modern dust supply from the Trade Winds (Fig. 1a) which is the third major sediment source in areas that are close to the coast, cannot be separated from the present day Harmattan dust in our grain-size analysis. Only additional petrographical evidence (abundance of chlorite, illite, and pale quartz) can distinguish the proper Trade Wind load from the Harmattan discharge. As a result, modern dust deposits of

ordinary Trade Wind origin are confined to a small sector 20° to more than 25° N offshore (Fig. 1b). The Trade Winds are also reflected in the sediments by the coastal upwelling activity that centres in this region³¹.

Deduction of wind speeds

Parkin³² developed the principles of a technique which relates the grain sizes of aeolo-marine dust deposits to the strength of the aeolian transport. We have developed this technique¹⁴ and applied it to our sets of grain-size data to estimate the regional patterns of both the present day and late Quaternary wind 'speeds' off North-west Africa.

Calculations here are based on several assumptions: for example that the different wind directions and their correlated dust qualities can be distinguished. Furthermore, we assume that the initial grain-size distribution $\delta V(l_0)/\delta D$ (where V = volume of dust per air unit; D = grain diameter; l_0 = zero transport distance) is proportional to $1/D$ (ref. 33). Wolter¹⁴ calculated possible examples and concluded that deviations from this function are not critical for the order of magnitude of the results. Finally, the grain-size spectrum in the air can be deduced from the grain-size distribution in the deep-sea sediment^{14,32}.

The wind-speed calculations begin with the mass-balance equation of aeolian-dust transport

$$u\delta V/\delta D - (u(\delta V/\delta D + \delta/\delta l(\delta V/\delta D)dl)) - w\delta V/\delta D = 0 \quad (1)$$

(where u is the horizontal wind 'speed', l is the dust-transport distance, and w is the dust particle terminal fall velocity). In

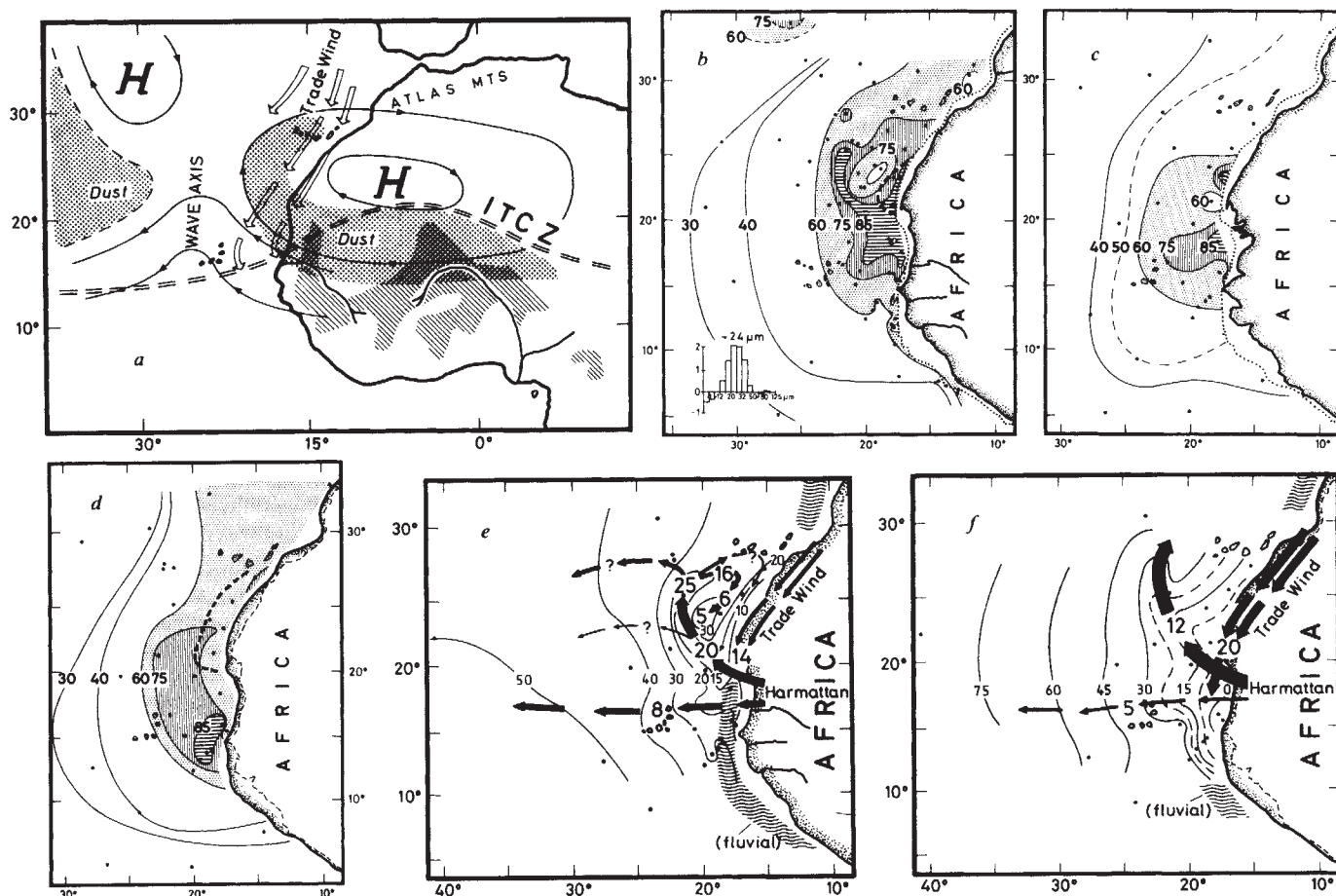


Fig. 1 a, Lower mid tropospheric flow patterns and underlying Trade Winds associated with an Easterly Wave and a Saharan air outbreak during July^{11,13}. Soil classification (oblique hatching)¹⁶ illustrates some source areas of lateritic trace particles. ITCZ, indicates the average summer position of the intertropical convergence zone at the surface. b, c, d, Distribution of the weighted terrigenous silt mode around 24 μ m (grain-size assemblage of Q-mode factor 2) in North Atlantic sediment samples indicating the dispersal pattern of proximal south Saharan/Sahelian dust. (b, Present; c, climatic optimum at ~6,000 yr BP; d, last glacial maximum at ~18,000 yr BP.) Area with dominant Trade Wind dust supply^{17,18,20} is surrounded by a broken line. e, f, Distribution patterns of $l/(uz)$ (equation (1)) (small numbers) and the major flow lines of north African dust. Wind speeds derived from equation (1) in $m s^{-1}$ (large numbers). Areas with fluvial sediment deposition shown by an excess of clay fraction (<6 μ m) are hatched. (e, Present; f, last glacial maximum at ~18,000 yr BP.) Width of arrows gives a rough estimate of dust quantities supplied (according to sedimentation rates²⁶).

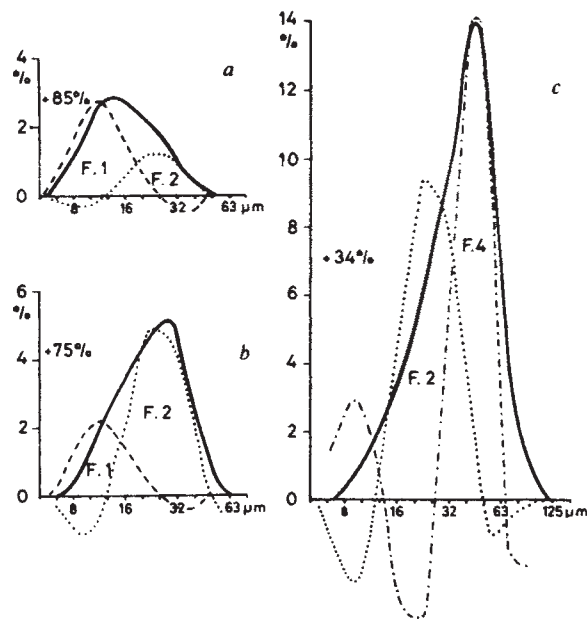


Fig. 2 Grain-size distributions of terrigenous silt classified into grain-size assemblages (Factors 1, 2, and 4) as defined by Q-mode analysis. Percentages (+) indicate the fraction $<6 \mu\text{m}$. Sample *a*, fluvial mud off Cape Verde (14°N , 19°W); sample *b*, Harmattan dust; sample *c*, synglacial proximal Trade and Harmattan Wind dusts (samples *b* and *c* off Mauritania, 18°N , 18°W).

equation (1) unidirectional dust input is balanced by unidirectional dust output and the gravitational dust fallout from the turbid volume of air. By integration, we can relate for each individual sample and for two arbitrary grain-size diameters, the gradient $\delta V/\delta D$ of the grain-size spectra in the air plume to the term $l/(uz)$:

$$l/(uz) = \frac{\ln((\delta V(D_1)/\delta D)/(\delta V(D_2)/\delta D) \cdot (D_1/D_2))}{k(D_2^2 - D_1^2)} \quad (2)$$

where z is the thickness of the turbid atmospheric dust layer which is 4 km on the average³⁴, and k equals the viscosity of the air.

Basically, any grain-size pair can be used in equation (2). A decrease of errors with increasing diameter, is suggested by error analysis¹⁴ using the redundancy of the information contained in the large data set now available. However, the grain diameters that occur on the extremely coarse tail ($>50 \mu\text{m}$) of the dust grain-size distribution have a high degree of statistical inaccuracy. The diameters $D_1 = 16 \mu\text{m}$ and $D_2 = 40 \mu\text{m}$ gave optimum accuracy.

The areal distribution of the $l/(uz)$ values (Fig. 1e), clearly reflects the pattern of proximal river supply and of dust distribution by the Harmattan and Trade Winds and therefore confirms the features that were found by Q-mode analysis (Fig. 1b). Modern dustwind 'speeds' estimated from equation (2) were characterized for two major streamline systems. They reach values of almost 20 m s^{-1} with a maximum of about 25 m s^{-1} in the right turn of the northern anticyclonic branch of the Harmattan, that is, in the dust vortex (Fig. 1e). A low 'speed' of only $\sim 8 \text{ m s}^{-1}$ occurs in the southern branch of the Harmattan, in the wake of the Easterly Waves. In addition, the 'speed' of the Trade Wind was estimated to be $\sim 14 \text{ m s}^{-1}$ 21° – 25°N offshore both from the sediments and from direct surface measurements³⁵.

Fossil record

The Q-mode analysis²⁷ and, to some extent, the dust-transport equation (2) was applied in an investigation of the variability of the depositional regimes that are represented by late Quaternary aeolo-marine dust sediments. The latter were from 32 deep-sea cores sampled in the subtropical area of the eastern Atlantic. They provide a continuous record of the atmospheric

circulation and extend beyond 50,000 yr BP (ref. 26). Two time slices were selected from this set of data to contrast the last two climatic extremes with today's climate. The chronostratigraphy is based on $\delta^{18}\text{O}$ and foraminiferal ecostratigraphic curves^{36,37} which were calibrated by ^{14}C dating up to $>40,000$ yr BP. The sources and modes of transport of the late Quaternary terrigenous sediments were inferred from their analogy with modern depositional patterns.

In a time slice of 6,000 yr BP which corresponds to the end of the last climatic optimum, the coarse proximal dust sediments (factor 2: $\sim 24 \mu\text{m}$) have been replaced largely by fine-grained material (complementary factor 1: $\sim 11 \mu\text{m}$; Fig. 1c). It is derived from long distance dust transport and, close to the continent, from an extensive fluvial supply of mud where an excess of clay is observed in the bimodal grain-size distributions^{18,26}. Its widespread and typically small scale lobate distribution means that the dust-transport equation (2) cannot be applied to this time slice. The deposits from both the Trade Wind dust and the coastal upwelling^{31,38} have disappeared almost completely, which indicates a considerably weaker meridional surface-wind component from NNE. However, the Harmattan-related dust outbursts continued to be active but were reduced, and the positions of their centres remained unchanged at $\sim 16^\circ$ – 20°N .

Sediments resembling any present-day river supply were absent from 10° – 12°N to north of 28°N during the last glacial maximum at 18,000 yr BP. Abundant sediments of the Harmattan-dust type then widely dominated the sedimentation pattern (Fig. 1d) but the coarse fallouts ($24\text{-}\mu\text{m}$ silt mode) hardly penetrated further west than they do today. In spite of the reduced sample coverage, the resemblance of the sickle-shaped dust pattern to the modern one is obvious. It parallels the modern course some 300 km further south. Adjoining the northeastern area, a fringe of even coarser grain sizes (factor 4: $\sim 45 \mu\text{m}$) extended up to 200 km off the coast 18° – 27°N . This sediment differs from the Harmattan load and contains dominant illite in the clay fraction, with unstained quartz north of 20°N (Fig. 1d). By analogy with modern deposits, this material may be regarded as aeolian dust and as related to an intensified glacial Trade Wind regime from NNE^{20,26}.

Applying the dust-transport equation (2) to the grain-size data from 18,000 yr BP (Fig. 1f), suggests that the zonal dustwind 'speeds' were considerably less than those of the present day (Fig. 1e). On the northern branch of the Harmattan, these decreased from 20 – 25 m s^{-1} to $\sim 12 \text{ m s}^{-1}$; on the southern branch in the wake of the Easterly Waves, they decreased from 8 to 5 m s^{-1} . On the other hand, the 18,000 yr BP Trade Winds with 'speeds' of 20 m s^{-1} seem to have been much stronger than at the present and this has been noted also for other parts of the world^{7,32,39}. These winds induced a vigorous coastal upwelling (as indicated by the associated high oceanic fertility^{31,40}) and they conveyed the fallout from the high altitude dust-loaded Harmattan-dust vortex further towards SSW. Trade Winds transported these particles according to their grain size over distances of up to several hundred kilometres.

However, the actual centre of the Harmattan dust outbursts from the African continent did not move southwards; its position during glacial times must have been similar to the present day, remaining constant at 15° – 20°N . This can be deduced from the persistently similar composition of the dust deposits south of 20°N , with dominant kaolinite in the clay fraction, and with reddish quartz, both of which are particularly characteristic of the Harmattan supply²⁰.

Conclusions

Our results on the three-dimensional dust transport systems have a direct bearing on the mechanisms of climatic change affecting the late Quaternary and present Saharan arid belt. The latitudinal position of the Harmattan-dust trajectory remained constant and this implies an on the long-term average constant summer position of the ITCZ and also of the adjacent upper-air high-pressure system over the western Sahara throughout

glacial and interglacial times. This stable pattern compares very well with the (summer and winter) position of the North Atlantic subtropical water mass as found in the planktonic foraminiferal record⁴¹. In the late Quaternary, both the northern boundary of this water mass and the associated west-wind drift zone, remained almost persistently stable west off Portugal. They hardly ever moved southwards, even during the glacial maximum with its extensive southward expansion of the North Atlantic polar water. Hence, it should not be the tropical but the Mediterranean and temperate zones which appear to be the climatic belts that are squeezed and expanded by the varying extension of the north-polar ice cap. Thus a mechanism of shifting climatic belts¹⁻⁴ can no longer explain the contrasting widespread glacial aridity and interglacial wetness all over North-west Africa, as recorded by fluctuating amounts of groundwater recharge⁴², fluvial sediment supply to the ocean and by the varying extent of dune fields on the continent⁷. Rather, the climatic changes are related to the varying strength of the tropical disturbances and the varying supply of water vapour from the equatorial Atlantic⁴³⁻⁴⁵. For example, decreased zonal winds are inferred from the weakened glacial

dust flow in the wake of the Easterly Waves. The variable strength of the meridional Trade Wind component seems to be a primary factor in the changing degree of wetness. This increased from almost zero during the Holocene climatic optimum to $>20 \text{ m s}^{-1}$ at 18,000 yr BP. It then had to compensate for the strong and increased meridional temperature gradients^{41,46}. Such a figure is consistent with the models of Manabe and Hahn⁴⁴ and Newell *et al.*⁴⁷ which postulate a glacial increase of meridional wind speeds by as much as 40% compared with the present, and with the record of the widespread (dormant) glacial seif dunes in the Sahel zone^{6,7,48}. Finally, the Trade Winds control the intensity of the oceanic upwelling and the correlated water-vapour and CO₂ budgets^{47,48}. Our results suggest that the mechanisms controlling the climate in North-west Africa during the present phase of the interglacial, are more closely related to that of the last glacial than to that of the Holocene climatic optimum.

We thank W. H. Berger, H. Flohn, and J. Schwarzacher for reviewing the manuscript and R. Stein and J. Mienert for computer and technical assistance. This work was supported by the Deutsche Forschungsgemeinschaft and by the CLIMAP project.

Received 25 March; accepted 30 July 1981.

1. Penck, A. *Sber preuss. Akad. Wiss.* **14**, 77-97 (1913).
2. Fairbridge, R. W. in *Problems in Palaeoclimatology* (ed. Nairn, A. E. M.) 356-360 (Wiley, London, 1964).
3. Rognon, P. & Williams, M. A. J. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **21**, 285-327 (1977).
4. Nicholson, S. E. & Flohn, H. *Climatic Change* **2**, 313-348 (1980).
5. Talbot, M. R. in *The Sahara and the Nile* (eds Williams, M. A. J. & Faure, H.) 37-62 (Balkema, Rotterdam, 1980).
6. Maley, J. thesis, Univ. Montpellier (1980).
7. Sarnthein, M. *Nature* **271**, 43-46 (1978).
8. Semmelhack, W. *Annls Hydrogr.* **62**, 273-277 (1934).
9. Jaenicke, R. & Schütz, L. *J. geophys. Res.* **83**, 3585-3599 (1978).
10. Prospero, J. M. & Carlson, T. N. *J. geophys. Res.* **77**, 5255-5265 (1972).
11. Carlson, T. & Prospero, J. *Rep. U.S. Geol. Survey Workshop (NCAR)*, Boulder, Colorado, 57-78 (1977).
12. Reed, R. J., Recker, E. E. & Norquist, D. C. *Mon. Weath. Rev.* **105**, 317-333 (1977).
13. Tetzlaff, G. & Wolter, W. *Palaeoecol. Afr.* **12**, 31-42 (1980).
14. Wolter, K. thesis, Univ. Hannover (1980).
15. Michel, P. *Mem. ORSTOM* **63**, 752 (1973).
16. Schmidt-Lorenz, R. *Mitt. Bund. Forsch. Anst. Forst-Holz* **1**-10 (1977).
17. Biscaye, P. E. *Bull. geol. Soc. Am.* **76**, 803-832 (1965).
18. Sarnthein, M., *et al.* in *The Geology of the Northwest African continental margin* (eds von Rad, U. *et al.*) (Springer, Heidelberg, in the press).
19. Sarnthein, M. & Koopmann, B. *Palaeoecol. Afr.* **12**, 239-253 (1980).
20. Sarnthein, M. *Met. Forsch. Erg. C* **31**, 49-51 (1979).
21. Radczewski, O. in *Recent Marine Sediments* (ed. Trask, P. D.) 496-502 (American Association of Petroleum Geologists, Tulsa, 1939).
22. Diester-Haass, L. *Quat. Res.* **6**, 299-314 (1976).
23. Milliman, J. D. *Deep-Sea Res.* **24**, 95-103 (1977).

24. Prospero, J. M., Giaccum, R. A. & Nees, R. T. *Nature* **289**, 570-572 (1981).
25. Kolla, V., Biscaye, P. & Hanley, A. *Quat. Res.* **11**, 261-277 (1976).
26. Koopmann, B. *Met. Forsch. Erg. C* **35** (in the press).
27. Kipp, N. G. *Geol. Soc. Am. Mem.* **145**, 3-42 (1976).
28. Egluff, J. *Bull. Am. Ass. petrol. Geol.* **56**, 694-706 (1972).
29. Embley, R. W., Rabinowitz, P. D. & Jacobi, R. D. *Earth planet. Sci. Lett.* **41**, 419-433 (1978).
30. Smayda, Th. J. *Mar. Geol.* **11**, 105-122 (1971).
31. Diester-Haass, L. in *Upwelling Ecosystems* (eds Boje, R. & Tomczak, M.) 261-280 (Springer, Heidelberg, 1978).
32. Parkin, D. W. *Proc. R. Soc. A* **377**, 73-100 (1974).
33. Schütz, L. & Jaenicke, R. *J. appl. Met.* **13**, 863-870 (1974).
34. Prospero, J. M., Savoie, D. L., Carlson, T. N. & Nees, R. T. in *Saharan Dust* (ed. Morales, C.) 171-186 (Wiley, Chichester, 1979).
35. Sarnthein, M. & Walger, E. *Geol. Rdsch.* **63**, 1065-1087 (1974).
36. Erlenkeuser, H., Koopmann, B. & Sarnthein, M. *Oceanol. Acta* (submitted).
37. Pflaumann, U. *Palaeoecol. Afr.* **12**, 191-212 (1980).
38. Petit-Maire, N. *Palaeoecol. Afr.* **12**, 365-377 (1980).
39. Molina-Cruz, A. *Quat. Res.* **8**, 324-338 (1977).
40. Müller, P. J. & Suess, E. *Deep-Sea Res.* **26**, 1347-1362 (1979).
41. Ruddiman, W. F. & McIntyre, A. *Geol. Soc. Am. Mem.* **145**, 111-146 (1976).
42. Sonntag, C. *et al. Palaeoecol. Afr.* **12**, 159-171 (1980).
43. Newell, R. E., Herman, G. F., Gould-Steward, S. & Tanaka, M. *Nature* **253**, 33-34 (1975).
44. Manabe, S. & Hahn, D. G. *J. geophys. Res.* **82**, 3889-3909 (1977).
45. Williams, J. in *Climatic Change* (ed. Gribbin, J.) 178-190 (Cambridge University Press, 1978).
46. CLIMAP Project Members. *Science* **191**, 1131-1137 (1976).
47. Newell, R. E., Navato, A. R. & Hsiung, J. *Pure appl. Geophys.* **116**, 351-371 (1978).
48. Henning, D. & Flohn, H. *Contr. Atmos. Phys.* **53**, 430-441 (1980).
49. Rona, P. A. *Deep-Sea Res.* **18**, 321-327 (1971).

Dispersion of α -like globin genes of the mouse to three different chromosomes

Aya Leder*, David Swan†, Frank Ruddle‡, Peter D'Eustachio‡ & Philip Leder*§

*Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, NIH, Bethesda, Maryland 20205, USA

†Laboratory of Cellular and Molecular Biology, National Cancer Institute, NIH, Bethesda, Maryland 20205, USA

‡Biology and Human Genetics Department, Yale University, New Haven, Connecticut 06520, USA

§Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA

The three active α -globin genes of the mouse are physically linked at a locus little more than 25 kilobases long. These genes form a part of an unexpectedly large and dispersed multigene family, including at least two pseudogenes located on two different chromosomes. This finding suggests that related gene sequences can be dispersed from their primary loci and that their dispersion may be a critical factor in the evolution of higher organisms.

ONE of the reaffirmed lessons of genetics is that genes move, that is, they migrate from place to place on and between chromosomes (for review see ref. 1). A rather newer lesson regarding genes in higher organisms is that they often consist of complex families of related sequences, with active genes often bracketing apparently inactive pseudogenes. The β -globin genes of mouse, rabbit, goat and man provide examples of this arrangement²⁻⁸.

One might have expected the α -globin genes of the mouse to conform to this model; there are two adult α genes (α -ad1 and α -ad2) and one embryonic gene^{2,9-11} as well as two closely-related pseudogenes^{2,12,13}. One of the pseudogenes has a novel structure in that it has lost its two intervening sequences in strict accordance with the GT/AG rule of RNA splicing^{12,13}. The other pseudogene is intact and still closely resembles its active, adult counterpart². In an attempt to understand the

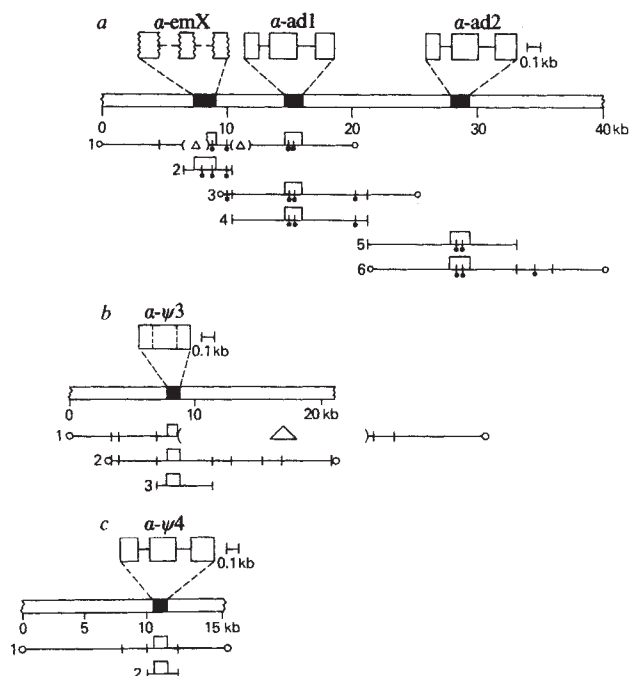


Fig. 1 Physical maps of the α -globin gene and pseudogene loci. *a*, Representation of a 40-kb long sequence encoding the α -emX, α -ad1 and α -ad2 genes. *b* Represents the α - ψ 3 locus, the 'splice-less' pseudogene, and *c* the α - ψ 4 locus. Filled boxes represent the location of each indicated gene; the enlarged diagram above each gene represents the distribution of the coding (open boxes) and intervening (single lines) sequences. The embryonic gene (α -emX) is shown with jagged borders along its coding sequence as it represents a putative rather than an established structure. This hypothetical representation is based on electron microscopic analysis and the expectation, by analogy to other active globin genes, that a second intervening sequence will be found (see Fig. 2). The clones used in mapping are indicated underneath, with the restriction sites for *Eco*RI (—|—) and *Bam*HI (—|—). The *Eco*RI fragments were cloned in the λ gtWES vector¹⁷ after isolation by preparative gel electrophoresis¹⁶. Established techniques were used for hybrid phage packaging³³ and clone identification³⁴. The overlapping clones were derived from a λ Charon 4A library formed by attaching *Eco*RI linker sequences to a partial *Hae*III digest of BALB/c mouse DNA by the method of Maniatis *et al.*¹⁸. The open circles at the ends of each library clone are synthetic *Eco*RI sites. The sequences indicated by Δ refer to deletions that have occurred in the hybrid phage as a consequence of cloning and phage propagation.

evolutionary origin of these genes and the mechanisms responsible for their amplification, we use recombinant DNA techniques to draw them into a coherent physical map. We describe here the cloning and physical arrangement of the adult and embryonic α -globin genes of the mouse and show that two α pseudogenes are dispersed away from the active locus from which they probably arose; they map on different chromosomes. While several mechanisms can account for the dispersion of these genes, the structure of the splice-less pseudogene is easily explained by imagining that it was conveyed to its new location via an RNA intermediate, possible a retrovirus¹⁴. The intact pseudogene may have been conveyed by more conventional transposition.

Physical map of the active α -globin gene locus

The BALB/c mouse produces two α -globins that appear in the adult organism, as well as a single embryonic α -globin that appears only in the fetus⁹. The two adult α -globins differ from one another at a single position, one having a serine, the other a threonine, at position 68 (ref. 15). We have previously cloned and determined the entire structure of the α -globin gene that encodes serine 68 (ref. 11). We have now used gene purification procedures¹⁶ to isolate and clone the *Eco*RI fragments cor-

responding to the second adult α gene and to the embryonic α gene. These genes were cloned in the vector λ gtWES using standard techniques^{16,17}. The two adult genes and the single embryonic gene were then linked to each other on a physical map by cloning overlapping fragments of DNA from phage λ libraries that incorporated partially digested restriction enzyme fragments of BALB/c mouse genomic DNA¹⁸.

The three genes were arranged into a physical map using restriction endonuclease digestion and fragment cross-hybridization. The map is shown in Fig. 1*a*; an α -embryonic sequence (α -emX) is closest to the 5'-end followed by α -ad1 and α -ad2. The active genes are located within a ~25-kilobase (kb) region. The embryonic α gene was located using cross-hybridization with a cloned cDNA fragment corresponding to the human embryonic α mRNA, as it does not share sufficient homology to allow cross-hybridization with the adult gene. The embryonic α gene was then identified within the cloned fragment using electron microscopy to visualize the R-loop created when the gene was annealed to mRNA derived from nucleated embryonic red cells (Fig. 2). (The embryonic gene neither formed R-loops nor cross-hybridized with adult globin mRNA in *in situ* hybridization analyses (data not shown).)

A comparison of the R-loop structures formed between the embryonic and adult genes and their respective mRNAs is shown in Fig. 2. The two genes show intervening sequences of different size, the adult gene having a classical structure (indicated in Fig. 2*b*) with two small intervening sequences; the embryonic gene has a larger intervening sequence of ~600 base pairs (bp) near its 5' end. A second intervening sequence cannot be visualized within the embryonic gene, but such an interruption, especially if it is less than 100 bases, may become apparent on sequencing the gene. The orientation of the genes within the fragments was determined by direct nucleotide sequencing in the case of α -ad2 (A.L., unpublished results) and by *in vitro* transcription of the gene in the case of α -emX (C. Talkington and A.L., unpublished results).

Efforts to extend the map of the active locus so as to locate the two α pseudogenes, α - ψ 3 and α - ψ 4, were unsuccessful. The pseudogene maps were determined by analysing partial restriction fragment libraries (see Fig. 1*b, c*), but did not overlap with any fragments located at the active α locus. Rather than extend the map further, we used the technique of chromosomal mapping in mouse $\times$ Chinese hamster hybrid cell lines¹⁹ to ensure that, at the very least, these genes would be located on chromosome 11, known to be the site of the adult and embryonic α genes²⁰. This hybrid cell technique has been used successfully for mapping several mouse and human genes²¹⁻²³.

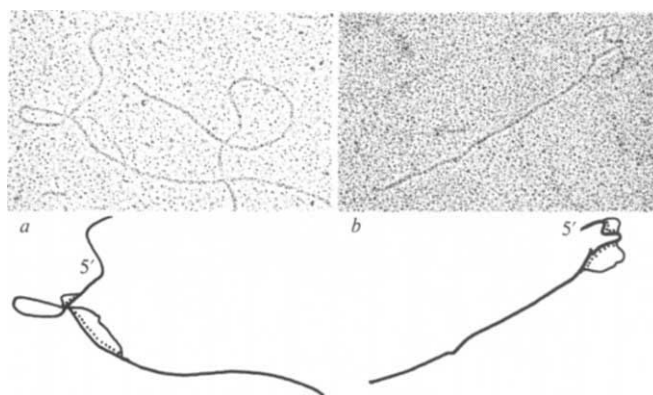


Fig. 2 Electron microscopic visualization of embryonic (*a*) and adult (*b*) α -globin genes. Genomic DNA fragments and mRNAs were purified and prepared for electron microscopy as previously described¹⁰. mRNAs were prepared from adult reticulocytes and from nucleated, 11-day-old embryonic yolk sac red cells. The diagrammatic representation of each R-loop is shown beneath each electron micrograph; the thick line represents double-stranded DNA, the thin line is the displaced DNA strand, and the dotted line indicates the RNA strand of an RNA/DNA hybrid.

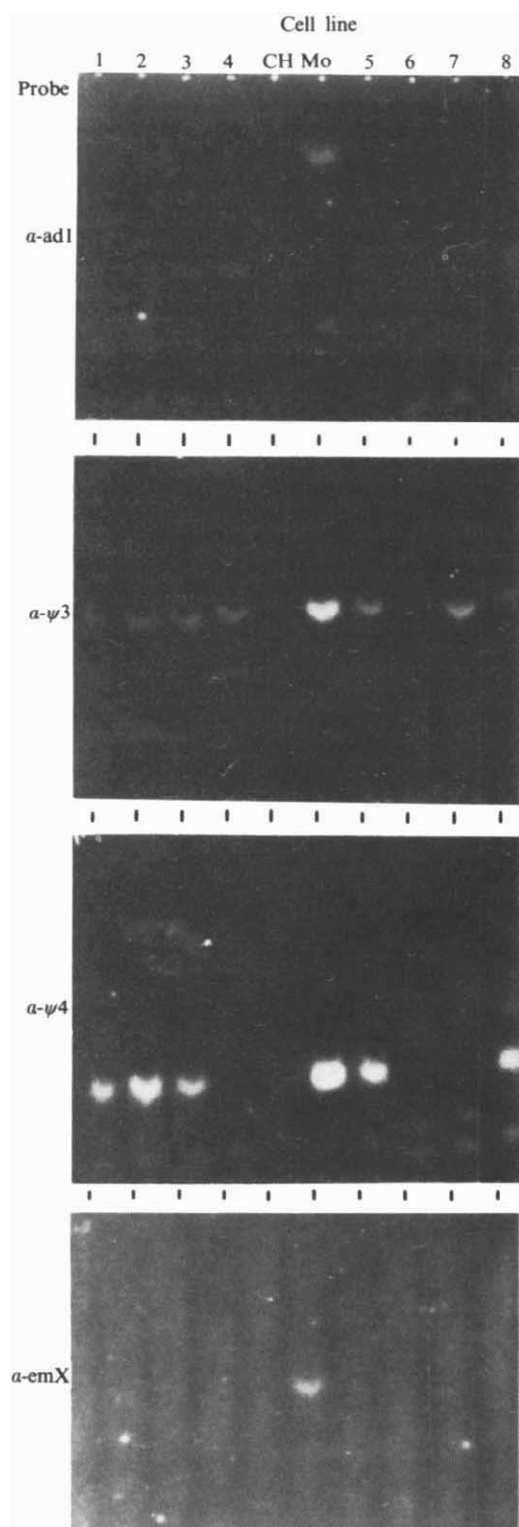


Fig. 3 Detection of α -globin-like gene sequences in mouse $\times$ Chinese hamster hybrid cell lines. Samples (30 μ g) of hamster (CH), mouse (Mo) and hybrid cell line (7A13-8, lane 1; 2A-1, lane 2; 2A-2, lane 3; 4A6-4, lane 4; BEM1-6, lane 5; 6D3Az, lane 6; Ecm4e, lane 7; and 4B31Az3, lane 8) DNA were digested with *Eco*RI, fractionated by agarose gel electrophoresis and transferred to diazotized paper as described elsewhere¹⁹. Fragments containing the α -like globin sequence were detected by hybridization to four different probes: α -ad1, mixed *Hha* fragments representing the 5' and 3' coding and flanking portions of the adult gene¹¹; α - ψ 3, the cloned *Eco*RI fragment shown in Fig. 1b; α - ψ 4, the cloned *Eco*RI fragment shown in Fig. 1c; and α -emX, the cloned *Eco*RI fragment shown in Fig. 1a.

α Genes and pseudogenes are located on different chromosomes

To determine the chromosomal location of the pseudogenes, we obtained hybrid cell lines between mouse and Chinese hamster that retained various numbers of mouse chromosomes¹⁹. The chromosomal content of each cell line was determined from cell samples taken for this analysis. The results are shown in Table 1. DNA samples isolated from nine mouse $\times$ Chinese hamster somatic hybrid lines were digested with the restriction endonuclease *Eco*RI and analysed using the *in situ* hybridization technique of Southern²⁴. DNA probes were made corresponding to α -ad1, α - ψ 3, α - ψ 4 and α -emX (see Fig. 3 legend). The results are shown in Figs 3 and 4 and are summarized at the bottom of Table 1. The adult and embryonic gene sequences only hybridize to DNA from the mouse, as none of the hybrid cell lines contain chromosome 11. However, the α - ψ 3 and α - ψ 4 genes hybridize to several of the cell lines, indicating that the two genes are on different chromosomes and that neither is on chromosome 11. These data are consistent with the location of the α - ψ 3 gene on chromosome 15 and the α - ψ 4 gene on chromosome 17. Note in Fig. 3, lane 7, that DNA from a cell line which contains only mouse chromosomes 14 and 15 is shown to contain only the α - ψ 3 gene. This gene is also present in DNA from a cell line having no mouse chromosome 14 (lane 1). Note also that Fig. 4, lane X, shows DNA from a cell line having only mouse chromosome 17 that contains the α - ψ 4 sequence alone.

What mechanism is responsible for moving the pseudogenes?

In view of the sequence homology between the active genes and pseudogenes (refs 2, 11–13 and Y. Nishioka, J. Maizel and P.L., unpublished), it is reasonable to imagine that the pseudogenes arose at the α locus by a gene duplication event and were, for a time, active at that locus²⁵. Their dispersion to other chromosomes could have occurred by various mechanisms, including transposition, chromosomal translocation, copy insertion or tetraploidization. Indeed, tetraploidization seems a plausible mechanism to account for the separation of the α - and β -globin genes which reside on a single chromosome in amphibia²⁶, but on separate chromosomes in higher genera. In the case of the mouse α pseudogenes, however, translocation and tetraploidization seem unlikely, given the isolated nature of these dispersed sequences. Copy insertion remains a feasible mechanism, although there is no precedent for this in vertebrates.

Given the above, transposition, the conveyance of a DNA fragment from one chromosomal location to another by a specific mobility element, can be considered a likely mechanism, especially in the case of the α - ψ 3 pseudogene. Several studies



Fig. 4 Detection of α - ψ 4 gene in a mouse $\times$ Chinese hamster hybrid cell line containing only mouse chromosome 17. The experiment was carried out as described in Fig. 3 legend. The lanes show DNA taken from the following sources; X, hybrid cell line R44.1; CH, Chinese hamster; MO, mouse. Negative controls used: α -ad1 and α -emX probes (not shown).

Table 1 Hybrid cell lines and controls tested for α -globin genes and pseudogenes

Mouse chromosome no.	Lane* 1 7A13-8	2 2A-1	3 2A-2	4 4A6-4	Chinese hamster	Mouse A/HeJ	5 BEM1-6	6 6D3Az	7 Ecm4e	8 4B31Az3	X R44.1
1		62	53	48	—	+	61				
2	93	85	90	79	—	+	103			12	
3			22		—	+	94				
4		38	43		—	+	97	14			
5	5		2		—	+					
6	48	50	69	79	—	+	197†				
7	52	69	69	79	—	+	13			85†	
8		27	33		—	+	87			6	
9	32	58	65		—	+	23				
10		35	51		—	+	27				
11					—	+					
12	36	81	73		—	+	74				
13	36	4	22		—	+	77			3	
14		62	82		—	+	103		100†	6	
15	68	77	90	154	—	+	174		100†	12	
16		73	80	86	—	+	87			12	
17	60	69	82		—	+	87			165†	100‡
18	15	19	55	1	—	+	55			59†	
19	36	69	82	107	—	+	126			50	
X		46	41	107	—	+	185†				
Probes					Reaction with probes						
α -ad1/2	—	—	—	—	—	+	—	—	—	—	—
α -emX	—	—	—	—	—	+	—	—	—	—	—
α - ψ 3	+	+	+	+	—	+	+	—	+	+	NT
α - ψ 4	+	+	+	—	—	+	+	—	—	+	+

Mouse chromosomes were identified in metaphase spreads prepared from hybrid cells, and subjected to combined trypsin–Giemsa–Hoechst staining techniques. The number shown is the mean number of copies of the chromosome per 100 cells. NT, not tested.

* See Fig. 3.

† Includes copies of the chromosome occurring in the form of translocations.

‡ Chromosome presence determined by thymidine kinase selection.

have shown that the murine retroviruses are formally analogous in structure to bacterial and yeast transposons (for review see ref. 27). Furthermore, genomic DNA that has been taken up by certain retroviruses seems to have lost its intervening sequences^{14,28}, a property also shown by the α - ψ 3 pseudogene. This similarity prompted Goff *et al.*¹⁴ to suggest that such a splice-less pseudogene might have been formed during its conveyance as part of a retrovirus sequence. Because α - ψ 3 retains sequence homology to the normal α -globin gene in the region 5' to the site of transcription initiation¹², it is unlikely to have been incorporated into a retrovirus as a cDNA copy of the globin mRNA, but rather as a genomic fragment. Such an integrated genomic fragment could then have been copied into viral RNA and subsequently acted on by the RNA splicing enzymes before being reintegrated into a germ cell chromosome as a part of the retrovirus sequence. In support of this notion, recent preliminary studies carried out by us, in collaboration with Kira Leuders and Edward Kuff, indicate that retrovirus-like A particle sequences²⁹ are encoded within a few thousand nucleotides of this gene.

What about the α - ψ 4 sequence? This gene has retained its intervening sequences and there is no compelling reason to assume that it has passed through a retrovirus during its evolu-

tion. Of course, transposition elements are likely to exist in other than retrovirus form and such sequences may have been responsible for the conveyance of this gene with its intervening sequences retained. If the transposition model is correct, it is possible that these sequences will be found in different configurations, conceivably on different chromosomes, in other strains of mice. Furthermore, it is possible that transposition of this gene could be followed in cultured somatic cells, just as the transposition of *copia* sequences was followed in somatic cells of *Drosophila*³⁰.

One further implication of the dispersion model is that there should be many more copies of α -globin sequences, possibly also dispersed throughout the mouse genome. To begin to test this possibility, we have used an adult α -globin cDNA clone³¹ to detect a family of more distantly related globin sequences that might be present in the mouse genome. We used a highly sensitive technique involving the two-dimensional separation of *Eco*RI-digested fragments of the mouse genome that could be detected by *in situ* hybridization³². This technique allows greater resolution of fragments and allows about 10 times as much DNA to be loaded onto the nitrocellulose strip, thus amplifying the signal obtained from weakly hybridizing gene sequences. Ordinarily, the globin cDNA probe detects four

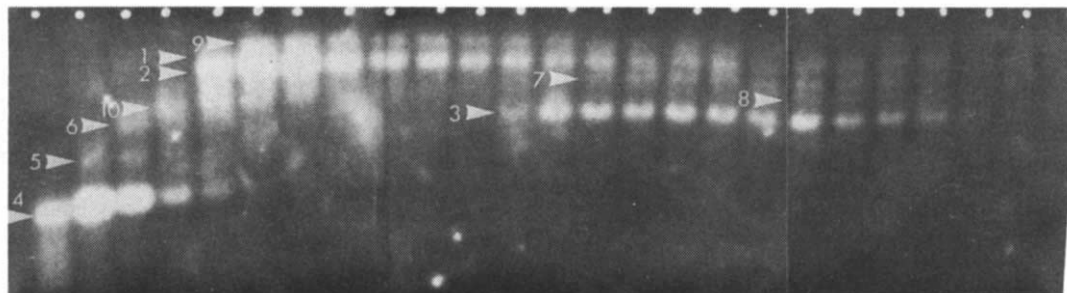


Fig. 5 Detection of cross-hybridizing *Eco*RI endonuclease fragments of genomic mouse DNA distributed by two-dimensional chromatography and electrophoresis. BALB/c mouse genomic DNA was digested with *Eco*RI restriction endonuclease, fragments were separated by RPC-5 column chromatography and subsequently by agarose gel electrophoresis (1% agarose) and transferred to nitrocellulose strips for *in situ* hybridization as described elsewhere³². The related fragments were detected using a mouse α -globin cDNA probe³¹ that was labelled to a specific activity of ~50 c.p.m. per pg of DNA. The filter was incubated in the presence of 6 $\times$ SSC at 65 °C and washed 5 times in 0.1 $\times$ SSC at 50 °C. The nitrocellulose sheet was sandwiched between Cronex image intensifying screens and exposed to Kodak XRP film for 4 days. Band numbers 1 and 2 represent fragments containing the two adult α -globin genes. Bands 3 and 4 represent fragments containing α - ψ 3 and α - ψ 4, respectively.

bands in a single dimension. These correspond to the two adult α genes and, hybridizing more faintly, the two α pseudogenes. The single embryonic gene is not detected. In contrast, using the more sensitive two-dimensional technique, 10 discrete bands are readily discerned (Fig. 5). (None of these bands correspond to β -globin gene-containing fragments¹⁶.) Evidently, many α -like globin sequences or fragments of sequences are present in the mouse genome, where their further characterization will provide a detailed test of the generality of the gene dispersion model.

An evolutionary role for pseudogenes

Pseudogenes appear to arise from active genes by some amplification event, possibly unequal crossing-over between homologous sequences. We further suggest that these additional gene copies remain active for a time and begin to drift slowly in sequence while under selective pressure. Should a new copy become redundant, for example, by further amplification, it would be free to accumulate mutations that render it inactive for the purposes of its original phenotype, that is, it would become a pseudogene. Such pseudogenes, unless they exert a deleterious effect on the organism, might be expected to remain at their original locus, accumulating mutational changes until they are no longer recognizable as descendants of their parental gene. However, certain other of these genes, given sufficient evolutionary time, will be recruited again as active gene sequences, not necessarily for their original purpose, but for some purpose

which they fulfill by chance. Furthermore, as these sequences already carry many of the elements required for gene expression, that is, promoter sites, splice signals, poly(A) addition sites, they possess a significant advantage in terms of their re-expression. From the results described here, we see that such genes can be scattered throughout the genome. The finding of many genomic fragments that cross-hybridize to globin genes further suggests that the genome may be a mosaic of conveyed pseudogenes at various stages of recruitment and decay, and that their formation and dispersion are critical factors in the evolution of the genomes of higher organisms. Indeed, the very recent description of isolated sea urchin histone pseudogenes (orphons) is likely to be an example of this phenomenon³⁵, differing only in that the isolated histone sequence originates from a very large tandemly arrayed multigene family and has yet to be mapped with respect to the active histone locus. The further analogy to gene dispersion in bacteria, yeast and *Drosophila*¹ provides a compelling precedent for such a general view.

We thank Terri Broderick for preparation of the manuscript and Barbara Norman and Marion Nau for their assistance. We also thank Dr Bernard Forget for providing the human embryonic α -like cDNA clone used in these studies; Dr Peter Lally for providing his unpublished data on the mapping of the α - ψ 3 and α - ψ 4 genes, which allowed us to reconcile an ambiguity in the mapping of the α - ψ 4 gene. This work was supported in part by grant GM09966 from NIGMS. P.D. was a special fellow of the Leukemia Society of America.

Received 7 May; accepted 30 July 1981.

- Calos, M. P. & Miller, J. H. *Cell* **20**, 579–595 (1980).
- Leder, P. *et al. Science* **209**, 1336–1342 (1980).
- Jahn, C. L. *et al. Cell* **21**, 159–168 (1980).
- Proudfoot, N. J., Shandler, M. H. M., Manley, J. L., Gefter, M. & Maniatis, T. *Science* **20**, 1329–1336 (1980).
- Lacy, E., Hardison, R. C., Quon, D. & Maniatis, T. *Cell* **18**, 1273–1283 (1979).
- Cleary, M. L., Haynes, J. R., Schon, F. A. & Lingrel, J. B. *Nucleic Acids Res.* **24**, 4791–4802 (1980).
- Jackson, I. J. & Williamson, R. *Br. J. Haemat.* **46**, 341–349 (1980).
- Bernards, R., Little, P. F., Annison, G., Williamson, R. & Flavell, R. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4827–4831 (1980).
- Russell, E. S. & McFarland, E. C. *Ann. N.Y. Acad. Sci.* **241**, 25–38 (1974).
- Leder, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 6187–6191 (1978).
- Nishioka, Y. & Leder, P. *Cell* **18**, 875–882 (1979).
- Nishioka, Y., Leder, A. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2806–2809 (1980).
- Vanin, E. F., Goldberg, G. I., Tucker, P. W. & Smithies, O. *Nature* **286**, 222–226 (1980).
- Goff, S. P., Gilboa, E., Witte, O. N. & Baltimore, D. *Cell* **22**, 777–785 (1980).
- Popp, R. A. *J. molec. Biol.* **27**, 9–16 (1967).
- Tilghman, S. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 4406–4410 (1977).
- Leder, P., Tiemeier, D. & Enquist, L. *Science* **196**, 175–177 (1977).
- Maniatis, T. *et al. Cell* **18**, 687–701 (1978).
- Swan, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 2735–2739 (1979).
- Russell, E. S. *Mouse News Lett.* **49**, 33 (1973).
- D'Eustachio, P., Pravtcheva, D., Marcu, K. & Ruddle, F. H. *J. exp. Med.* **151**, 1545–1550 (1980).
- D'Eustachio, P., Bothwell, A. L. M., Takaro, T. K., Baltimore, D. & Ruddle, F. H. *J. exp. Med.* **153**, 793–800 (1981).
- Gusella, J. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 2829–2833 (1980).
- Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
- Miyata, T. & Yasunaga, T. *Proc. natn. Acad. Sci. U.S.A.* **78**, 450–453 (1981).
- Jeffreys, A. J. *et al. Cell* **21**, 555–564 (1980).
- Bishop, J. M. A. *Rev. Biochem.* **47**, 35–88 (1978).
- De Feo, D. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Kuff, E. L., Smith, L. A. & Leuders, K. K. *Molec. Cell Biol.* **1**, 216–227 (1981).
- Potter, S. S., Brorin, W. J., Dunsinuir, P. & Rubin, G. *Cell* **17**, 415–427 (1979).
- Rougeon, F. & Mach, B. *Gene* **1**, 229–239 (1977).
- Tiemeier, D. C. *et al. Cell* **14**, 237–245 (1978).
- Hohn, B. & Murray, K. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3259–3263 (1977).
- Benton, W. D. & Davis, R. W. *Science* **196**, 180–182 (1977).
- Childs, G., Maxson, R., Cohn, R. H. & Kedes, L. *Cell* **23**, 651–664 (1981).

LETTERS TO NATURE

Synchronous extreme spectral variability of BE Ursae Majoris

Bruce Margon

Astronomy Department, University of Washington, Seattle, Washington 98195, USA

Ronald A. Downes & Jonathan I. Katz

Astronomy Department, University of California, Los Angeles, California 90024, USA

BE Ursae Majoris is a 15-mag variable star of period 2.29 days, with a sinusoidal light curve of ~ 1 mag amplitude¹. Recently, Ferguson *et al.*² have drawn attention to this object, also known as PG1155+492, because of its UV excess and unusual spectrum, which they observed to vary between absorption and strong emission on a time scale of months. Here we report observations which demonstrate that the spectrum of BE UMa is dramatically variable on a time scale of hours and days, in synchrony with the 2.29-day periodic light variations. We interpret the system as a detached binary consisting of an extremely hot ($T \sim 10^5$ K) white dwarf and a cool M star, whose upper layers are irradiated by the intense ionizing flux of the hot companion. This unique system should prove invaluable for measurements of the mass, luminosity and bolometric correction of the hottest white dwarfs.

Our spectrophotometric observations were made on three consecutive nights in March 1981 and four consecutive nights in April 1981. The March data were obtained with 30 Å spectral resolution in the 3,400–6,800 Å range, using the intensified image dissector scanner at the 2.1 m reflector of the Kitt Peak National Observatory. The April data consist of 10 Å resolution scans in the 3,700–6,800 Å band, using the Robinson-Wampler image tube scanner at the 1.5-m Mt Lemmon telescope. Our data may be summarized by noting that one of three basic spectral states can describe a given observation: (1) a blue continuum with strong, narrow Balmer, He I, He II and C III λ 4,650 emission; (2) a featureless, UV-excess continuum; or (3) a blue continuum with strong absorption lines of He II λ 4,686 and a strong, but not obviously identifiable, absorption feature at $4,050 \pm 15$ Å. Ferguson *et al.*² have seen this system in basically similar spectral states, in observations separated by months. To our surprise we have on several occasions observed the system switch between these states on a time scale of hours.

Figure 1 shows an example of some of the more dramatic spectral changes, with two spectra obtained 29 h apart (approximately one-half of the system period) displayed. In this brief interval the object has switched completely from the absorption to the emission state. Unfortunately a hardware malfunction in the data acquisition computer prevented us from converting these observations from instrumental counts to absolute fluxes; however, the two spectra were obtained with the identical

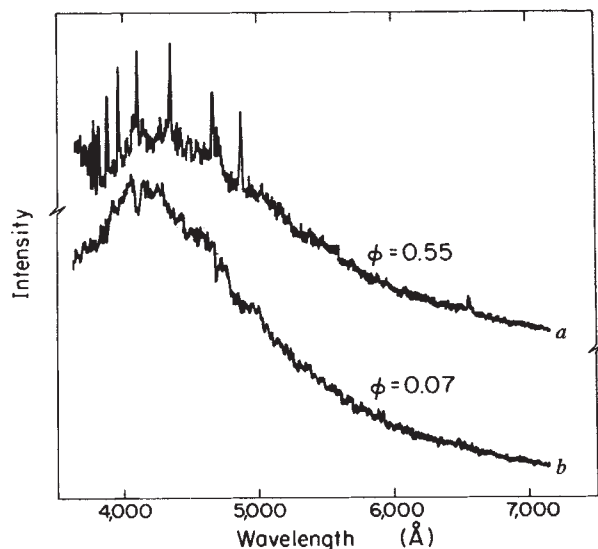


Fig. 1 Spectra of BE UMa obtained on consecutive nights, 9 March (b) and 10 March (a) 1981. The relative intensity scale is linear, with the zero-point at $\sim 7,000$ Å for both spectra. The photometric phase of each of the observations is shown, using the period of ref. 1, but assigning phase 0 to minimum light.

instrumental configuration and thus the slopes can be compared directly. There is no dramatic colour change associated with the spectral transition; the system may become slightly less UV in the emission state. Perhaps even more surprising, the emission lines occur near the phase of photometric maximum. This is in complete contrast to the cataclysmic variables³, whose spectra bear at least superficial resemblance to BE UMa in the emission state⁴. In the cataclysmics, featureless or absorption spectra are normally associated with the photometric outbursts and emission spectra with the quiescent periods³.

The spectral line strengths in Fig. 1 are of particular interest. The very prominent absorption line near 4,050 Å (equivalent width 3 Å) confirms the observation of this feature by Ferguson *et al.*². These authors speculated that it may be spurious in their data, partly because such a prominent feature has no obvious identification. We have seen this feature on several nights, and in some cases (see Fig. 1) it is the strongest feature in the spectrum. We can offer no new clues as to the origin. One of the more peculiar spectral features noted by Ferguson *et al.*² is that the C III $\lambda 4,650$ line appeared as the strongest emission line in the entire spectrum, exceeding even the Balmer lines and He II $\lambda 4,686$. Although we have one observation where this is also true, most of our spectra show $\lambda 4,650$ at ~ 0.8 the strength of $\sim 4,686$. The surprising strength of this C III feature could be due to Bowen fluorescence pumping², which has been observed to be responsible for the feature in the X-ray-irradiated atmosphere of HZ Her⁵. However, Ferguson *et al.* were unable to detect the O III $\lambda 3,444$ emission line, which should then be prominent. We have also searched for this feature without success in one particularly well exposed UV spectrum and can set upper limits on its intensity of 20% of that of C III.

Figure 2 shows an interesting example of the spectral variability of BE UMa on a time scale of hours. Three spectra fortuitously obtained through the photometric minimum on 11 March, with spacings of 5 and 3 h, have caught the system in transition from the emission to the continuous state. The entire transition can apparently occur on a time scale which is short compared with one 2.29-day period, but longer than appropriate for the eclipse of a compact star.

We have measured or set upper limits on the equivalent widths of the Balmer, He II $\lambda 4,686$ and C III $\lambda 4,650$ emission on 26 spectra obtained over the seven nights of observation, and find that there is a striking correlation of these quantities with the photometric phase of BE UMa. Figure 3 shows these data

folded with the 2.291067-day period determined by Kurochkin¹; note that, contrary to his phase convention, we set phase zero at minimum light. Although our observations are incomplete between phases 0.1 and 0.4, it can readily be seen that the emission strengths are sharply modulated by the period, with maximum strength emission repeatedly occurring near photometric maximum, and the featureless and/or absorption spectra occurring near minimum. We have several examples of spectra obtained within 1 h of minimum light that definitely show weak emission lines; it thus seems likely that at least on some cycles the emission never entirely disappears.

The relatively long-period, sinusoidal light curve¹ and prolonged transition between spectral states (Fig. 2) all argue that the observed variations in BE UMa are not the eclipses of a binary system. We consider instead models where the ionizing flux from a hot degenerate star heats a nearby companion, analogous to the situation observed in HZ Her/Her X-1. We identify the spectrum near minimum light as due to the photosphere of a hot white dwarf. It is implausible that the white dwarf's surface is heated by an invisible companion, for it intercepts only 10^{-6} of the companion's radiation, and a companion 10^6 more luminous than the white dwarf would be detectable in some region of the spectrum.

More promising are models where the white dwarf's intense ionizing flux heats the facing photosphere of a cool red dwarf companion. Minimum light of BE UMa then corresponds to inferior conjunction of this cool star, and the visual spectrum is dominated by the white dwarf. Maximum light corresponds to observation of the irradiated face of the cool companion, and the strong emission spectrum is due to reprocessed radiation from the white dwarf. A consistent model, where the cool star is not easily detectable in the visual region, may be obtained with a main sequence companion of $0.4 M_{\odot}$, approximately an M2V star. The observed period and an assumed white dwarf mass of $0.7 M_{\odot}$ then imply that the cool star intercepts 10^{-3} of the white dwarf's radiation. To explain the observed 1-mag photometric modulation, the re-radiated power must exceed the visible radiation from the white dwarf. Thus, depending on the reprocessing efficiency (which is probably not very different from unity), the bolometric correction (BC) for the white dwarf must approximately cancel the 10^{-3} dilution of ionizing flux, implying BC ~ -7.5 mag. Model atmospheres with white dwarf surface gravities^{6,7} yield this BC for $T \sim 10^5$ K. We thus infer that BE UMa contains one of the hottest known white dwarfs, consistent with the observed very steep UV spectrum². The (uncertain) distance of such a system would be of order 150 pc.

Our model explains both the sharp, intense emission spectrum and its pattern of variability. The reprocessing occurs chiefly in emission lines because the ionizing UV radiation is absorbed high in the cool star's atmosphere, a situation resembling the soft

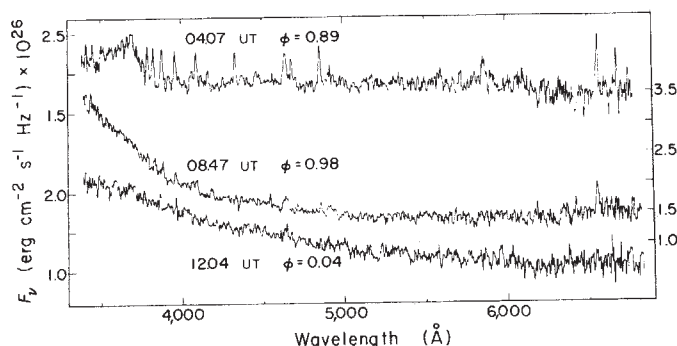


Fig. 2 Three spectra observed on 11 March 1981, during a transition from the emission to the absorption state. The data have been converted to absolute fluxes by observation of spectrophotometric standard stars. The universal time and photometric phase of each of the observations, which are typically of 40 min duration, are shown near each spectrum. The right-hand intensity scale applies to the middle spectrum.

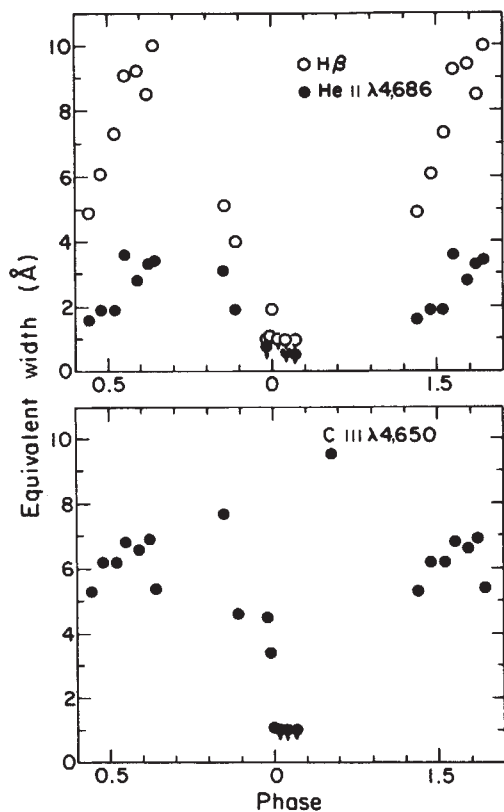


Fig. 3 The variation of emission line equivalent widths of BE UMa with photometric phase; slightly more than one cycle is shown for clarity. Observational uncertainties are $\sim \pm 1$ Å in equivalent width.

X-ray-heated atmosphere of Sco X-1 rather than the hard X-ray-heated HZ Her. The emission line intensities vary continuously with the aspect of the red dwarf. If the inclination is not very close to 90° , as is likely on purely statistical grounds, then some emission is expected throughout the orbital cycle, as is indeed observed. The red dwarf underfills its Roche lobe by a factor of ~ 5 in radius, so there is no reason to expect mass transfer. Thus the dissimilarities of BE UMa from cataclysmic variables² (for example, lack of flickering, narrow emission lines, unusual UV spectrum) are readily understood. One may wonder about the probability of one of the hottest known white dwarfs occurring in a close binary system. However, observational selection is surely an important factor, as it is exactly such systems that call attention to themselves through spectacular variable heating effects.

Our analysis implies several testable predictions. Detailed modelling of the UV and near IR data² must be compatible with the presence of the hot white dwarf and the M dwarf. We predict that the emission line spectrum will show periodic radial velocity variation of amplitude $\geq 100 \sin i \text{ km s}^{-1}$, with velocity extrema near photometric quadrature. The white dwarf absorption spectrum should show an anti-phased velocity variation of about half this amount, with the amplitude more sensitive to the uncertain mass of the M star. Sufficiently sensitive radial velocity observations may detect the white dwarf's gravitational redshift through a difference in the systemic velocities of the absorption and emission lines, thus adding BE UMa to the very select group of white dwarfs with measured masses. Independent (but uncertain) mass estimates may also become available through combining the observed radial velocity data with limits on the inclination derived from modelling the light curve and pattern of emission line equivalent width variations.

The evolutionary status of BE UMa presents several interesting questions. The white dwarf has presumably recently lost considerable mass to reach its current state, yet there is no overt evidence for this mass in the form of nebulosity². If the mass was

captured by the cool companion, the main sequence mass-radius relationship may not apply to that star, raising the possibility that the object is oversized for its temperature, and requiring a somewhat less extreme bolometric correction for the white dwarf. If the cool star is more massive than we have supposed, subsequent evolution could produce a cataclysmic variable. Systems whose cool star is more luminous than the present case would not be readily recognized as peculiar. BE UMa may thus be a recently formed member of that elusive class of objects, the cataclysmic variable progenitors.

Thus, we find BE UMa to be a unique hybrid of previously studied close binary systems with degenerate companions. Unlike the cataclysmic variables, there is probably no significant mass transfer, thus allowing eventual detailed modelling of the system. Heating effects are far more prominent than in previously studied hot white dwarf/red dwarf close pairs such as the extreme UV source Feige 24 (refs 8–10), which has a 4.2-day period. The ratio of ionizing-to-optical flux in BE UMa is similar to that in HZ Her/Her X-1, but in that system the heating occurs deep in the photosphere, leading to only subtle phase dependence of the spectrum rather than the very large amplitude effects observed here. It seems likely that BE UMa will figure prominently in future studies of the curious properties of stellar photospheres irradiated from above by a copious ionizing flux, as well as in the derivation of the mass, luminosity and bolometric correction of the most extreme end of the white dwarf luminosity function.

We thank D. Ferguson and J. Liebert for sharing their data before publication, E. Böhm-Vitense and P. Szkody for useful discussions, and the NSF for financial support. B.M. and R.A.D. were visiting astronomers at the Kitt Peak National Observatory, which is operated by AURA for NSF; J.I.K. and B.M. are Alfred P. Sloan Foundation Research Fellows.

Received 18 May; accepted 20 July 1981.

1. Kurochkin, N. E. *Perem. Zvezdy* **15**, 77–83 (1964), **18**, 85–90 (1971).
2. Ferguson, D., Liebert, J., Green, R., McGraw, J. T. & Spinrad, H. *Astrophys. J.* (submitted).
3. Warner, B. *IAU Symp.* **73**, 85–140 (1976).
4. Ferguson, D. H., Liebert, J., Green, R. F., McGraw, J. T. & Spinrad, H. *Bull. Am. astr. Soc.* **12**, 851 (1980).
5. Margon, B. & Cohen, J. G. *Astrophys. J. Lett.* **222**, L33–L36 (1978).
6. Wesemael, F., Auer, L. H., Van Horn, H. M. & Savedoff, M. P. *Astrophys. J. Suppl.* **43**, 159–303 (1980).
7. Wesemael, F. *Astrophys. J. Suppl.* **45**, 177–257 (1981).
8. Margon, B., Lampton, M., Bowyer, S., Stern, R. & Paresce, F. *Astrophys. J. Lett.* **210**, L79–L82 (1976).
9. Liebert, J. & Margon, B. *Astrophys. J.* **216**, 18–22 (1977).
10. Thorstensen, J. R., Charles, P. A., Margon, B. & Bowyer, S. *Astrophys. J.* **223**, 260–265 (1978).

An X-ray pulsar in SNR G109.1–1.0

G. G. Fahlman* & P. C. Gregory†

Departments of *Geophysics and Astronomy and †Physics, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5

The unusual extended X-ray source G109.1–1.0 (ref. 1), discovered with the Einstein Observatory², looks like a supernova remnant (SNR) and is of particular interest because it contains a point source 1E2259+586, from which a long curving arc of emission appears to extend, joining smoothly with the outer edge of the diffuse emission. We have now analysed the photon arrival times for the point source. (The point source was previously^{1,3} designated as GF2259+586; here we will use the 1E designation adopted by Tuohy and Garmire⁴.) We conclude that the point source is an X-ray pulsar.

The data were obtained with the Image Proportional Counter² between 23 h 26 min 57 s UT, 7 July and 1 h 37 min 20 s, 8 July 1980. Within this 2-h time interval the data are contained in three distinct segments, as shown in Fig. 1. The count rate illustrated includes an estimated 7% contribution from the extended source and the general background. This time series

was analysed at the University of British Columbia using a straightforward period-folding approach. A systematic search for periods in the range $0.1 \text{ s} \leq P \leq 200 \text{ s}$ quickly revealed a strongly modulated signal with a period of $3.4890 \pm 0.0002 \text{ s}$ in the last data segment of Fig. 1. Figure 2a shows this light curve. The period and degree of modulation are quite consistent with known X-ray pulsars⁵. Figure 2b,c shows the light curves obtained by folding the two shorter data segments at the same period. A fixed phase point was used for all three light curves, so the phase shift in Fig. 2 is real. Light curves obtained in different partitions of the third segments did not reveal any obvious phase shift from the average presented in Fig. 2a.

We have also analysed the discovery observation of 17 December 1979. These data have a much poorer signal-to-noise ratio because the source is located considerably off-centre in the field. Nevertheless, the same periodic modulation is clearly present. Thus the pulsation is not a transient phenomenon.

The short period detected is consistent with the hypothesis that a rotating neutron star is the underlying source of the pulsating flux. There are three possible mechanisms for producing the X-ray luminosity: (1) conversion of rotational energy as in radio pulsars, (2) accretion from the ambient interstellar medium, and (3) accretion from a close companion. The first possibility seems implausible because the period is much longer than is the case for the only known examples of this phenomenon, the Crab and Vela pulsars, and there is no radio or optical counterpart apparent (see below). As for the second possibility, if the association with the diffuse X-ray source is accepted, then the ambient density would be far too low to power an X-ray pulsar by accretion. The possibility of the point source being either a foreground or background object superimposed on the extended source is improbable in view of its appearance. This leaves the third possibility. Available evidence suggests that most, if not all, the X-ray pulsars (except the Crab) are in binary systems⁵ and this is probably the case for 1E2259+586. Indeed, it is possible that the phase shift noted above is due to orbital motion. However, we cannot place firm constraints on possible orbits until further data become available.

An accurate position for 1E2259+586, obtained with the High Resolution Imager², is $\alpha = 22 \text{ h } 59 \text{ min } 3.4 \text{ s}$, $\delta = 58^\circ 36' 38''$ (1950.0). No optical counterpart is visible at this position on either of the relevant Palomar Sky Survey prints nor on a deeper red plate obtained by D. Crampton at the prime focus of the Canada-France-Hawaii Telescope. A conservative limit on the apparent magnitude of the source is $V \geq 21$. Our radio observations, to be discussed in detail elsewhere, place an

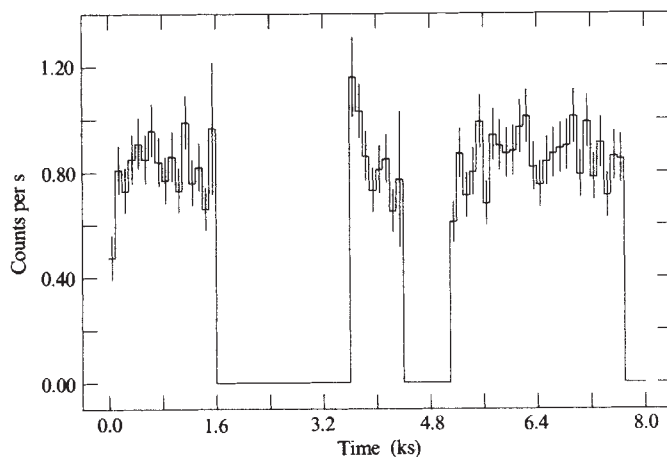


Fig. 1 The X-ray flux of the central source 1E2259+586 on 7–8 July 1980 shown in 100-s bins. The error bars are based on the counting statistics and the ordinate unit is the internal proportional counter count rate uncorrected for background and diffuse source counts.

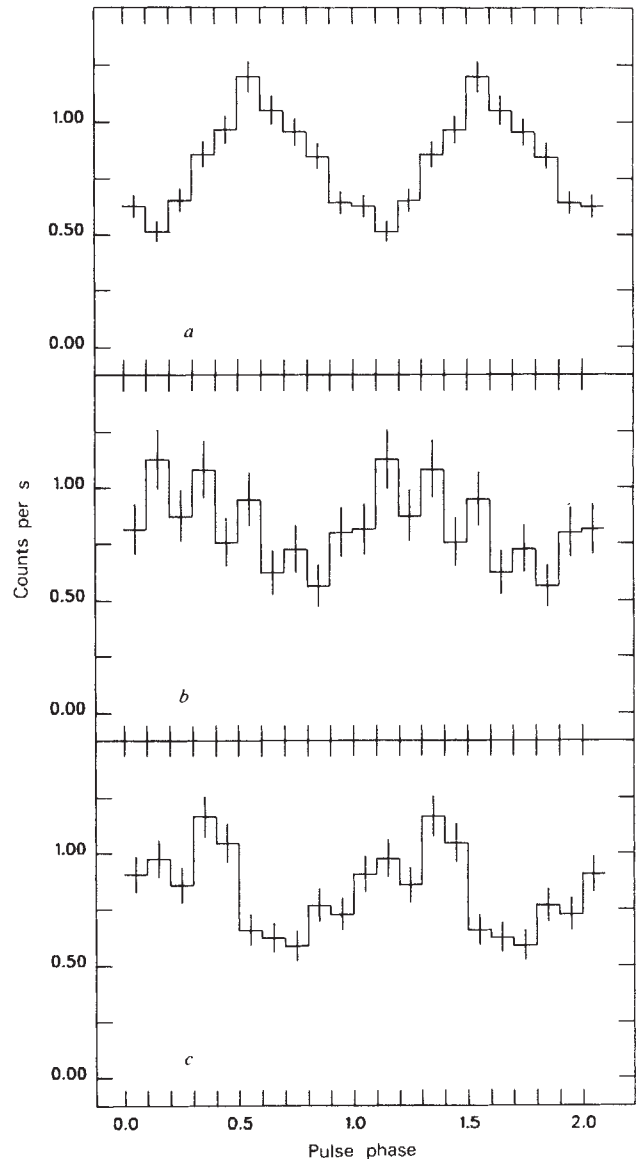


Fig. 2 Light curves obtained by folding the data at the pulsar period of 3.4890 s. One cycle contains 10 phase bins and is shown twice; *a* is the data in the last segment of Fig. 1, *b* is for the middle segment and *c* applies to the first segment. A common phase reference time was used for all three light curves. The error bars are the square root of total counts in each phase bin.

upper limit of 0.5 mJy at 20 cm for any compact radio source coincident with 1E2259+586. (There is a compact radio source ~ 2 arc min south-west of the X-ray source but this is probably a background radio galaxy.)

We know of only four other compact X-ray sources associated with SNRs: the Crab and Vela pulsars, the binary system SS433 (refs 4, 6) and a newly reported source in RCW103 (ref. 7) whose nature is unclear. 1E2259+586 is the only known accretion-powered X-ray pulsar in this group and probably joins SS433 as the second binary system. This binary system clearly differs from SS433. If we adopt a distance of 4.0 kpc and an extinction of $A_V = 4.0$ mag (see ref. 1), then the lack of a visible counterpart ($V \geq 21$) puts a lower limit of $M_V \geq 4$ mag on the companion to 1E2259+586. This rules out a supergiant or massive main-sequence star and suggests that the source is a compact low-mass binary^{8,9}. This hypothesis is appealing because the short orbital periods expected in such a system may readily explain the short-term phase shift in the light curves shown in Fig. 2. On the other hand, it is not clear whether a binary system could evolve into a compact configuration within

the time interval set by inferred age for the SNR. A definitive orbital analysis is required to resolve this issue.

The X-ray image suggests a continuing interaction between the point source and the SNR shell. The point source is an X-ray pulsar and, almost certainly, a member of a binary system. This fact will be important in understanding the origin of the interaction and concomitantly, the evolution of the remnant itself.

We thank F. D. Seward for assistance with the Einstein data. This work is supported by operating grants from the Natural Science and Engineering Council of Canada.

Received 9 June, accepted 20 July 1981.

1. Gregory, P. C. & Fahlman, G. G. *Nature* **287**, 805 (1980).
2. Giacconi, R. et al. *Astrophys. J.* **230**, 540 (1979).
3. Gregory, P. C. & Fahlman, G. G. *Proc. int. Meet. on SS433*, Rome (in the press).
4. Seward, F., Grindlay, J., Seaquist, E. & Gilmore, W. *Nature* **287**, 806 (1980).
5. Rappaport, S. & Joss, P. C. Preprint CSR-HEA-80-24 (MIT Center for Space Research, 1980).
6. Crampton, D., Cowley, A. P. & Hutchings, J. B. *Astrophys. J. Lett.* **230**, L131 (1980).
7. Tuohy, I. & Garmire, G. *Astrophys. J. Lett.* **239**, L107 (1980).
8. Joss, P. C. & Rappaport, S. *Astr. Astrophys.* **71**, 217 (1979).
9. Li, F. K., Joss, P. C., McClintock, J. E., Rappaport, S. & Wright, E. L. *Astrophys. J.* **240**, 628 (1980).

The X-ray cataclysmic variable 1E0643.0–1648

Jeremy Bailey*, J. H. Hough*, A. Schwarzenberg-Czerny†, Ian Gatley‡, R. F. Jameson§ & D. A. Hanes||

* Hatfield Polytechnic Observatory, Bayfordbury House, Hertford SG13 8LD, UK

† Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA, UK

‡ United Kingdom Infrared Telescope Unit, 900 Leilani Street, Hilo, Hawaii, USA

§ Astronomy Department, University of Leicester, Leicester LE1 7RH, UK

|| Anglo-Australian Observatory, PO Box 296, Epping, New South Wales 2121, Australia

We have used a new simultaneous IR/optical high-speed photometer on the UK IR telescope to study the recently discovered X-ray cataclysmic variable 1E0643.0–1648 (refs 1, 2). The light curve of this object obtained by the American Association of Variable Star Observers (AAVSO) (T. Chlebowski, personal communication) has shown it to be a dwarf nova with a recurrence time scale of 15 days. We report here that our photometry, obtained during the decline from an outburst, shows slow flickering, with the optical and IR light curves correlated with no delay. We have also obtained moderate resolution spectra of the star using the Anglo-Australian telescope.

Photometry was obtained on the nights of 8, 9 and 10 January 1981 using the 3.8-m UK IR telescope at Mauna Kea. The photometer uses a dichroic mirror to separate the IR and optical beams. The IR is fed to a solid nitrogen-cooled InSb detector. The optical beam is further divided by a partially reflecting mirror, a portion of the light being fed to a cooled GaAs photomultiplier tube while the remainder is available for guiding. Background subtraction for both IR and optical channels is provided by the UKIRT *f*/35 chopping secondary mirror operating at ~ 5 Hz. Data acquisition is by a microcomputer system which records continuous data runs using integration times which can be as short as the chop cycle. The light curves are displayed in real time on a visual display unit (VDU) and stored on disk for subsequent analysis.

The observations of 1E0643.0–1648 used integration times of 5 s and the filters were V (0.55 μm) and J (1.25 μm). The light curves obtained on 9 and 10 January are shown in Fig. 1. The data have been binned in 25-s intervals, as this star shows little variability on very short time scales. A shorter run of data on 8 January showed the star at $V \sim 11.5$. Thus the star faded by 0.8 mag at V over 2 days but showed little change at J. The AAVSO light curve shows that the star was declining from an outburst at

this time. The reddening as the star declines probably indicates a significant contribution from the red secondary star at J.

As previously reported^{1,3}, the star shows flickering on time scales of minutes in the V light curve, and we find similar variability with smaller amplitude at J. To study this variability further the data have been analysed using autocorrelation, cross-correlation and power spectrum techniques. The autocorrelation (ACF) and cross-correlation (CCF) functions, calculated according to the definitions given by Stockman and Sargent⁴, are plotted in Fig. 2. The mean and slope were subtracted from each data set before analysis. No correction has been made for the zero delay spikes in the ACFs. The power spectra of the light curves show no significant periodicities down to periods of 10 s. We can thus make the following comments.

(1) The e-folding time of the ACFs is typically 3 to 4 min. This is a measure of the characteristic time scale of individual flares. The time scale of X-ray flickering also seems to be a few minutes².

(2) IR and optical flickering are clearly correlated with no significant delay between them, as indicated by the CCFs and their similarity to the ACFs. This sets a limit on the physical separation of the emitting regions responsible for the optical and IR flickering of $\sim v\Delta t$ where v is the velocity of propagation of the disturbance responsible for the flickering and Δt here is ~ 10 s. If we take the conventional model of dwarf nova flickering, in which the flickering arises in a bright spot on the outer edge of the accretion disk due to inhomogeneities in the accretion stream, then the relevant value of v is a stream velocity of the order of a few hundred km s^{-1} . This gives a maximum separation of a few $\text{cm} \times 10^8$, which is less than the probable size of the bright spot and would imply that the IR and optical flickering regions are essentially the same. However, another possibility is that the optical/IR flickering arises from reprocessing of the X rays by the disk or secondary star. This is particularly plausible for this system in view of the similar time scales observed for optical and X-ray flickering. In this case the maximum separation would be $c\Delta t$ which is comparable with the size of the binary system.

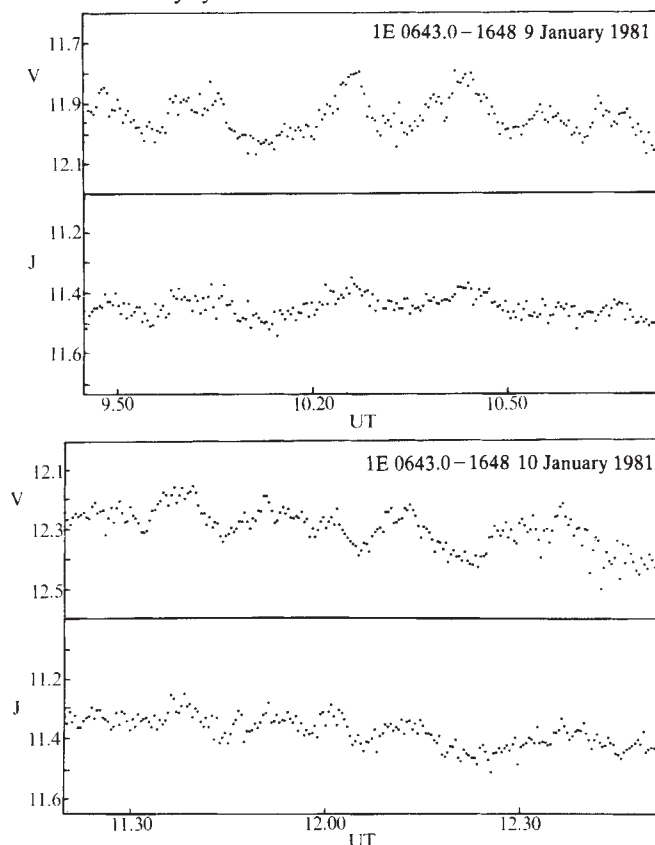
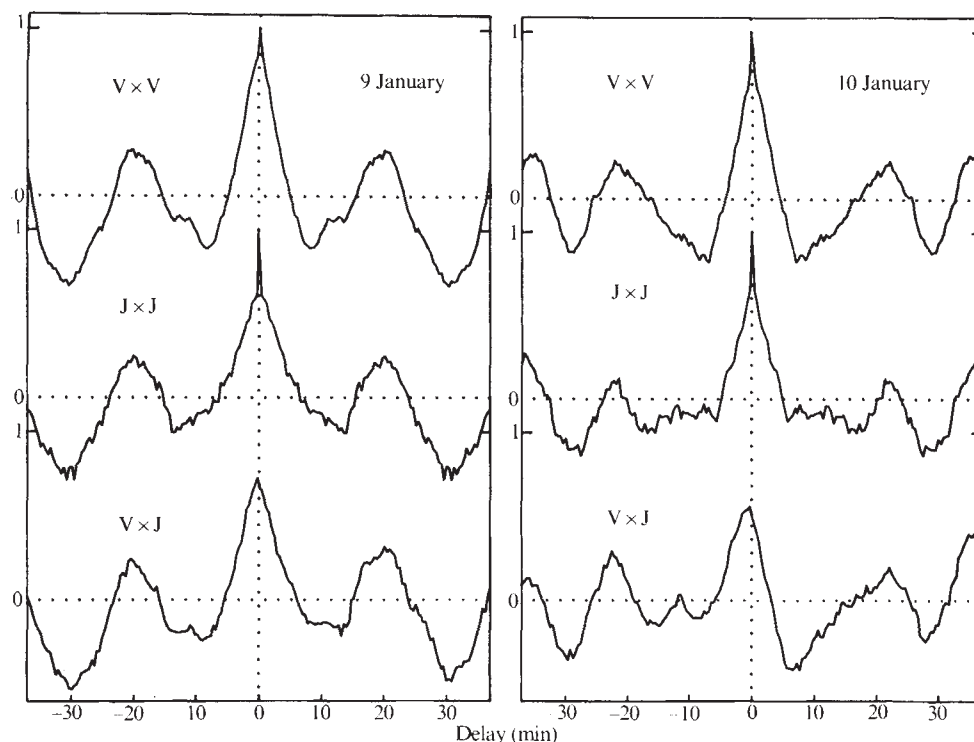


Fig. 1 V and J light curves of 1E0643.0–1648 observed on 9 and 10 January 1981.

Fig. 2 Autocorrelation function of V data ($V \times V$), autocorrelation function of J data ($J \times J$) and cross-correlation function of V and J data ($V \times J$) obtained from the 9 and 10 January light curves.



(3) The amplitude of flickering at J is smaller than that at V. To estimate this effect quantitatively the data were filtered to retain only the frequencies associated with the flickering, and the regression line of J on V was fitted to the data sets. The slope of this line is a measure of the colour associated with the flickering. It was found that on both nights the flickering was $0.9 (\pm 0.1)$ mag bluer in V-J than the system as a whole. Note that the colour of the flickering need not be the same as that of the emitting region responsible for the flickering. For example, if flickering is due to varying temperature in a source emitting as a black body, the spectrum of the source is given by the Planck function, but the spectrum of the flickering is given by the temperature derivative of the Planck function. By this assumption the observed flickering colours correspond to a temperature of 7,000–10,000 K.

(4) The secondary peak at a delay of 20 min in the ACFs and CCFs indicates the typical repetition time scale of the larger flares. Although a peak at 20 min is present in all the power

spectra, we cannot consider this as evidence for a significant periodicity in view of the relatively few cycles observed.

(5) For dwarf novae there is an approximate relationship between the rate of decline in brightness following outburst, and the orbital period^{5,6}. From our observations and those of Warner³ it seems that the decline rate is ~ 0.4 mag per day, which would imply an orbital period of ~ 6 h. The recent report⁷ of a spectroscopic period of 5.2 h is consistent with this value. Thus we cannot comment on the form of the binary light curve as our observations cover substantially less than one cycle per night.

Two narrow slit spectra of 1E0643.0–1648 were obtained centred on 26.71 December 1980 in cloudy conditions with the 3.9-m Anglo-Australian telescope, using the RGO spectrograph with the Boksenberg image photon counting system⁸. The spectra cover the region 3,950–4,450 Å with 0.8 Å resolution and the integration times were 1,000 s.

Figure 3 shows the sum of the two spectra. The general

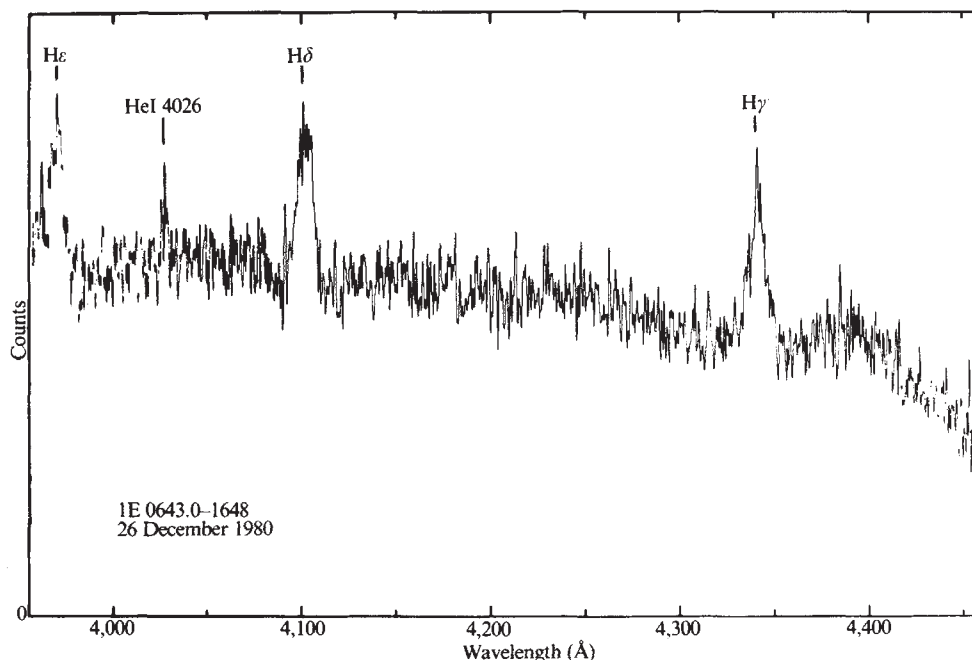


Fig. 3 Spectrum of 1E0643.0–1648 observed on 26 December 1980. This is the sum of two spectra, each of 1,000 s integration.

appearance is consistent with that reported elsewhere^{1,2}. The presence of strong hydrogen emission lines and weak He I lines makes the spectrum appear similar to those of other cataclysmic variables. The equivalent widths of the lines determined from the spectrum in Fig. 3 are: He I, 3 Å; He I 4026, 0.6 Å; Hδ, 5.1 Å and Hγ, 6.5 Å. We find no evidence in our spectra for Zeeman features as suggested by Chlebowski *et al.*². There is no evidence for doubling of the Balmer emission lines as is seen in high inclination systems such as Z Cha (ref. 9), OY Car (ref. 10) and WZ Sge (ref. 11 and C. C. Brunt, personal communication). This may be taken as evidence for a relatively low orbital inclination. After Rayne and Whelan⁹, the width of the line can be used to obtain a lower limit to the white dwarf mass if we assume that the lines arise in an accretion disk. From Fig. 3 we estimate the full width to be $1,300 \text{ km s}^{-1}$ at zero intensity of the Hγ and Hδ lines. Interpreting this width as due to the keplerian velocity of the innermost parts of the disk, we obtain the following limits using Webbink's¹² white dwarf mass-radius relation. For $i = 60^\circ, 45^\circ, 30^\circ$ and 15° , $(M_1/M_\odot)_{\min} = 0.08, 0.12, 0.2$ and 0.5 , respectively. Except for very low inclinations these are much lower than the masses normally inferred for the white dwarf components of cataclysmic variables, which suggests either that the inclination is very low, or that the disk does not extend to the surface of the white dwarf. The latter case might result from the presence of a magnetic field which disrupts the inner disk.

We thank the staff at UKIRT and AAT for their assistance. We particularly thank David Beattie for his work in commissioning the simultaneous optical photometer, Mark Bateson for assistance with some of the UKIRT observations, and J. A. J. Whelan for discussions and assistance with the AAT observations. We also thank T. Chlebowski for information and a preprint, PATT for the award of AAT and UKIRT time and the SRC for financial and travel support.

Received 11 May; accepted 21 July 1981.

- Chlebowski, T., Halpern, J. P. & Steiner, J. E. *IAU Circ. No. 3529* (1980).
- Chlebowski, T., Halpern, J. P. & Steiner, J. E. *Astrophys. J. Lett.* **247**, L35 (1981).
- Warner, B. *IAU Circ. No. 3560* (1981).
- Stockman, H. S. & Sargent, T. A. *Astrophys. J.* **227**, 197 (1979).
- Bailey, J. *J. Br. astr. Ass.* **86**, 30 (1975).
- Bailey, J. thesis, Univ. Sussex (1978).
- Hutchings, J. B., Cowley, A. P. & Crampton, D. *IAU Circ. No. 3585* (1981).
- Boksenberg, A. *Proc. ESO/CERN Conf. Auxiliary Instrumentation for Large Telescopes*, Geneva, 295 (1972).
- Rayne, M. W. & Whelan, J. A. J. *Mon. Not. R. astr. Soc.* **196**, 73 (1981).
- Bailey, J. & Ward, M. *Mon. Not. R. astr. Soc.* **194**, 17P (1981).
- Krzeminski, W. & Kraft, R. P. *Astrophys. J.* **140**, 921 (1964).
- Webbink, R. F. *Mon. Not. astr. Soc.* **171**, 555 (1975).

Heterogeneous photocatalytic oxidation of aromatic compounds on TiO₂

Masamichi Fujihira, Yoshiharu Satoh & Tetsuo Osa

Pharmaceutical Institute, Tohoku University, Aobayama, Sendai 980, Japan

The photoelectrochemistry at semiconductor electrodes has been used to design systems for heterogeneous photosynthesis¹⁻⁷ and photocatalysis⁸⁻¹⁰. There has been, however, no attempt at the organic synthesis of highly valuable compounds from cheap raw material except for a very low yield of amino acids¹¹ produced from methane-ammonia-water. If such products were to be formed in one step by heterogeneous photocatalytic reactions, this would be a highly efficient means for the economical utilization of solar energy in comparison with hydrogen production, even though the reactions are driven in a spontaneous ($\Delta G^\circ < 0$) direction by the light energy.

We tried to make positive use of the unfavourable 'short circuiting' of the heterogeneous photocatalysts in which the product from anodic reaction site can be reduced at the neighbouring reducing site. The photosynthesis of H₂O₂ on TiO₂ powders is one example of this type of reaction, where e^- and h^+ represent photoinduced electron and hole, respectively, in

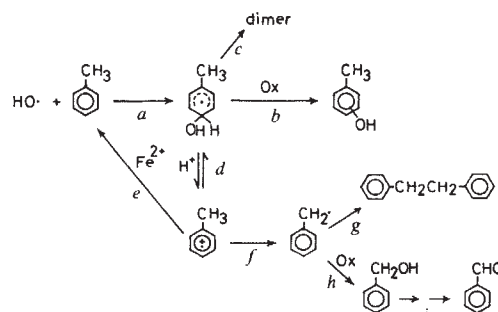
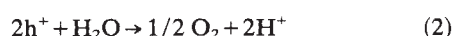


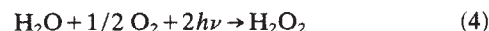
Fig. 1 Heterogeneous photocatalytic oxidation of toluene in various conditions.

the conduction and valence bands of the heterogeneous photocatalyst.



and for overall reaction ((1) × 2 + (2) + (3)).

By combining the principles of photochemistry at semiconductor electrodes with the Fenton reaction we show here how this reaction can be achieved.



Both the oxygen dissolved in water and that produced from water by the photogenerated holes (equation (2)) at the anodic reaction site is involved in the reduction to H₂O₂ by the excited electrons from the conduction band of a TiO₂ particle (equation (3)). Similar photosynthesis of H₂O₂ has been studied using ZnO (refs 12-16), CdS, CdSe, and other semiconductors¹⁷ in the presence or absence¹⁶ of organic reductants (such as alcohol, amide and acid). In the latter, however, the anodic reaction is not oxidation of water, but it causes corrosion of the semiconductor catalysts or oxidation of organic additives. Thus, oxygen evolution from water on the stable semiconductors such as TiO₂, SnO₂, and SrTiO₃ is a preferred photoanodic reaction.

The photosynthesis of hydrogen peroxide itself is probably important as the light energy is stored as chemical energy by the overall reaction (equation (4)) in a nonspontaneous direction ($\Delta G^\circ = 27.9 \text{ kcal mol}^{-1}$)¹⁸. But H₂O₂ is unstable^{13,18} especially in the presence of some metal ions and under illumination by UV light¹⁷. Consequently, we can use H₂O₂ as a starting material for further reaction *in situ* to produce more valuable compounds. For example we adopted a well known Fenton reaction¹⁹ in which H₂O₂ decomposes into hydroxyl radical and hydroxide anion by the reductive cleavage with ferrous ion²⁰:



the HO· formed is very reactive and is used for hydroxylation of aromatic ring and side chain-oxidation of alkylbenzenes (Fig. 1)²¹⁻²³.

We succeeded in the photo-Fenton type reaction utilizing H₂O₂ formed on TiO₂ powders from the photosynthetic reduction of oxygen which was either dissolved in air or was formed by photoanodic decomposition of water on the same TiO₂ photocatalyst. Benzene, toluene, and acetophenone were tried as the organic substrates. All the products expected from Fenton reaction were obtained, although some had already been found to be reaction intermediates for the oxidative photocatalytic decomposition of hydrocarbons to CO₂ (ref. 10).

TiO₂ photocatalysts (Aerosil P-25 and 99.99% anatase from Rare Metallic Co.) were used without further treatment. Each photocatalyst (1 g), in a mixture of doubly distilled water (50 or 100 ml) and an organic substrate (5 ml), was irradiated by a 500-W Xe-lamp or a 500-W high-pressure mercury arc lamp in a 100 ml Pyrex glass round-bottomed flask with a condenser cooled with running water. To keep the solution saturated with

air and to maintain homogeneous mixing of the suspension, the solution was stirred by a magnetic stirrer. If necessary, nitrogen was bubbled for deaeration. After photolysis, the solution was made acidified by HCl, then extracted with ether, and finally concentrated under a reduced pressure. A Shimadzu GC-4CM gas chromatograph was used to analyse the products. The amount of a gaseous product, CO₂, was determined¹⁰.

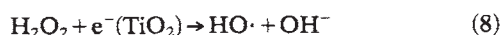
The results for organic substrate-TiO₂-H₂O systems without any other additives are listed in Table 1. In air, every aromatic compound was hydroxylated to give the corresponding phenol. Other than the hydroxylated compounds, biphenyl from benzene, benzaldehyde from toluene, and phenol from acetophenone were obtained. From toluene or acetophenone, another product which may be attributable to a dimer as biphenyl was also detected. When the benzene solution was deaerated before illumination by bubbling nitrogen for 30 min and sealed by nitrogen overflow during illumination, biphenyl was obtained as a main product with a trace amount of phenol. When nitrogen was bubbled throughout the irradiation, no product was detected from benzene and toluene. When the toluene solution was stirred vigorously for 10 h without irradiation, no product was detected. These results indicate that oxygen and illumination are necessary for the oxidation of organic substrates. If oxygen is removed completely, the only possible cathodic reaction will be reduction of proton to form hydrogen. But the low efficiency of this reaction^{1,24} on TiO₂ is why the absence of oxygen results in no oxidation of aromatic compounds. The result obtained by illumination under nitrogen sealing preceded by deaeration is due to the lower concentration of oxygen rather than its absence. Because no catalyst for efficient production of HO· was added, we consider the following reactions are responsible for HO· formation



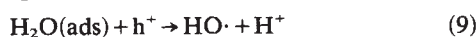
(the Haber-Weiss reaction)^{10,25}



(photodissociation with UV light, $\lambda \leq 365 \text{ nm}$)²⁶



(reductive cleavage at TiO₂ surface)



(intermediate of anodic reaction on TiO₂)²⁷

where reaction of equation (8) is a new proposal because the photosynthesis of H₂O₂ on TiO₂ has been studied in liquid phase but hitherto with negative results^{13,18}.

The effects of adding acid (H₂SO₄), oxidant (Cu²⁺), and catalyst for HO· radical formation (Fe²⁺) are shown in Table 2. Adding H₂SO₄ only, increases the total yield drastically, this could be due to the suppression of CO₂ formation (Table 2),

Table 1 Photocatalytic oxidation of aromatic compounds on TiO₂ powders*

Aromatic† compound	Conditions	Time of illumination (h)	Products	mg (μmol)
Benzene	Under air	12‡	Phenol	22.0(234)
			Biphenyl	4.8(31)
	After deaeration Under N ₂	10‡	Phenol	0.4(4)
			Biphenyl	2.1(14)
Toluene	Under air	2§	—	—
			—	—
	Under N ₂ Stirred for 10 h without illumination	2§	Benzaldehyde	1.1(10)
			Cresol (o : m : p)	1.4(13) (44 : 40 : 16)
Acetophenone	Under air	10‡	—	—
			—	—
	Under N ₂ Stirred for 10 h without illumination	0	—	—
			—	—
Acetophenone	Under air	10‡	Hydroxy AP (o : m : p)	11.4(84) (49 : 33 : 18)
			Phenol	3.5(37)
	Under N ₂ Stirred for 10 h without illumination	0	Dimer	~1.0(~4)
			—	—

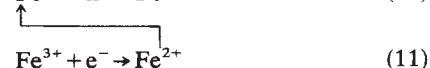
* Aerosil P-25. † Initial composition of solution: aromatic compound 5g, TiO₂ 1g, H₂O 50 ml, except toluene (5g, 1g, 100 ml). ‡ A 500 W Xe lamp. § A 500 W Hg lamp.

Table 2 The effect of additives on the product distribution of heterogeneous photocatalytic oxidation of toluene in TiO₂-H₂O in air*

Additives	None	H ₂ SO ₄ (0.05 M)	H ₂ SO ₄ (0.05 M) + CuSO ₄ (0.05 M)	H ₂ SO ₄ (0.005 M) + CuSO ₄ (0.2 M)	H ₂ SO ₄ (0.005 M) + FeSO ₄ (0.02 M)
Products (μmol)					
Benzaldehyde	11.8	234.3	646.4	111.3	30.8
Benzyl alcohol	—	trace	188.1	—	—
o-Cresol	32.7	—	80.7	208.2	3.7
m-Cresol	14.0	—	7.6	67.5	—
p-Cresol	9.9	—		89.7	—
Bibenzyl	—	—	70.8	—	—
CO ₂	(~100)	trace	†	†	†
Total products (μmol)‡	68.4	234.3	1064.4	476.7	34.5
Quantum efficiency (%)‡	0.3	1.1	5.0	2.2	0.16

* Illuminated for 2 h by a 500 W Hg lamp, initial composition of solution: toluene 5g, TiO₂ (99.99% anatase from Rare Metallic Co.) 1g, H₂O 100 ml. † Expected to be a trace, but not measured. ‡ Calculated on the basis of the amount of the starting material transformed, that is 1 mol of bibenzyl counted as 2 mol of products while other products counted as only 1 mol but only if CO₂ were excluded.

although the Haber-Weiss reactions as potential sources of the HO· radical in the absence of metal impurities have been predicted²⁵ for high and low pH. Evolution of CO₂ might be interpreted as the photocatalytic oxidative decarboxylation of carboxylates^{9,10} which may result from the further oxidation of the intermediates such as benzaldehyde and mucondialdehyde^{28,29}. If the pH of the solution is too low, an intermediate formed by the HO· attack on the aromatic ring³⁰, hydroxycyclohexadienyl radical, will be changed promptly to a cation radical (route *d* in Fig. 1)²³ and side chain-oxidation products will be the main products (routes *f*, *g*, *h*, and *i*). Presence of an oxidant and an increase in pH usually leads to hydroxylation of aromatic ring (route *b*) predominating over side chain-oxidation in the Fenton reaction²¹⁻²³. These predictions agree well with the observed results in Table 2. It is particularly surprising that lowering the pH and adding Cu²⁺ increased the quantum yield by up to 5%, which is about 16 times more efficient than would happen in an additive-free system. In the system with an Fe²⁺ catalyst for HO· formation, the yield became rather low which could be due to the 'short circuiting' on TiO₂ by the reversible redox reaction:



This was not the case for the Cu²⁺ reduction, in fact the reduction of oxygen prevailed³¹ over the reduction of Cu²⁺ at low pH.

We conclude that the combination of the photoelectrochemical production of H₂O₂ at semiconductor electrodes with the Fenton reaction enabled us to succeed in the heterogeneous photocatalytic oxidation of various aromatic hydrocarbons. By changing the solution conditions, that is by controlling pH and/or adding appropriate catalysts (Cu²⁺ ion in the present system), both the unstable product H₂O₂ and the reactive intermediate HO· could be utilized more efficiently to synthesize more stable and valuable compounds from aromatic hydrocarbons in high quantum yields (1–5%) compared with an additive-free system (0.3%). The appropriate solution conditions enabled selective formation of cresol (77%) or benzaldehyde (61%) from toluene. The present result suggests that the unfavourable 'short circuiting' character of the heterogeneous photocatalytic reactions can be utilized positively in other systems if the proper combination of semiconductor material and reactants and the solution conditions is selected.

Received 23 April; accepted 30 June 1981.

- Bard, A. J. *J. Photochem.* **10**, 59–75 (1979).
- Fujishima, A. & Honda, K. *Nature* **238**, 37–38 (1972).
- Bulatov, A. V. & Khidkel, M. L. *Izv. Akad. Nauk SSSR, Ser. Khim.* 1902–1903 (1976).
- Schrauzer, G. N. & Guth, T. D. *J. Am. chem. Soc.* **99**, 7189–7193 (1977).

5. Kawai, T. & Sakata, T. *Nature* **286**, 474–476 (1980); *JCS Chem. Commun.* 694–695 (1980).
6. Halmann, M. *Nature* **275**, 115–116 (1978).
7. Inoue, T., Fujishima, A., Konishi, S. & Honda, K. *Nature* **277**, 637–638 (1979).
8. Frank, S. N. & Bard, A. J. *J. Am. chem. Soc.* **99**, 303–304 (1977); *J. phys. Chem.* **81**, 1484–1488 (1977).
9. Kraeutler, B. & Bard, A. J. *J. Am. chem. Soc.* **100**, 2239–2240, 5985–5992 (1978).
10. Izumi, I., Dunn, W. W., Wilbourn, K. O., Fan, F. F. & Bard, A. J. *J. phys. Chem.* **84**, 3207–3210 (1980).
11. Reiche, H. & Bard, A. J. *J. Am. chem. Soc.* **101**, 3127–3128 (1979).
12. Baur, E. & Neuweiler, C. *Helv. chim. Acta* **10**, 901–907 (1927).
13. Markham, M. C. & Laidler, K. *J. phys. Chem.* **57**, 363–369 (1953).
14. Rubin, T. R., Calvert, J. G., Rankin, G. T. & MacNevin, W. M. *J. Am. chem. Soc.* **75**, 2850–2853 (1953).
15. Freund, T. & Gomes, W. P. *Catal. Rev.* **3**, 1–36 (1969).
16. Dixon, D. R. & Healy, T. W. *Aust. J. Chem.* **24**, 1193–1204 (1971).
17. Stephens, R. E., Ke, B. & Trivich, D. *J. phys. Chem.* **59**, 966–969 (1955).
18. Clechet, P., Martelet, C., Martin, J. R. & Olier, R. *Elect. chim. Acta* **24**, 457–461 (1979).
19. Walling, C. *Acc. Chem. Res.* **8**, 125–131 (1975).
20. Haber, F. & Weiss, J. J. *Proc. R. Soc. A* **147**, 332–351 (1934).
21. Smith, J. R. L. & Norman, R. O. C. *J. chem. Soc. B* 1013–1018 (1969).
22. Walling, C. & Johnson, R. A. *J. Am. chem. Soc.* **97**, 363–367 (1975).
23. Eberhardt, M. K. & Martines, M. I. *J. phys. Chem.* **79**, 1917–1920 (1975).
24. Ohnishi, T., Nakato, Y. & Tsubomura, H. *Ber. Bunsenges. Phys. Chem.* **79**, 523–525 (1975).
25. Weinstein, J. & Bielski, B. H. *J. Am. chem. Soc.* **101**, 58–62 (1979).
26. Gibson, J. F., Ingram, D. J. E., Symons, M. C. R. & Townsend, M. G. *Trans. Faraday Soc.* **53**, 914–920 (1957).
27. Archer, M. D. *J. appl. Electrochem.* **5**, 17–38 (1975).
28. Daniels, M., Scholes, G. & Weiss, J. J. *J. chem. Soc.* 832–834 (1956).
29. Loeff, I. & Stein, G. *J. chem. Soc.* 2623–2633 (1963).
30. Dorfman, L. M., Taub, I. A. & Bühler, R. E. *J. chem. Phys.* **36**, 3051–3061 (1962).
31. Reiche, H., Dunn, W. W. & Bard, A. J. *J. phys. Chem.* **83**, 2248–2251 (1979).

A wave theory for sandwaves in shelf-seas

F. D. C. Hammond & A. D. Heathershaw

Institute of Oceanographic Sciences,
Crossway, Taunton, Somerset TA1 2DW, UK

Sea-waves are understood more easily if they are considered to move with constant velocity and without changing shape. A convenient mathematical technique is to imagine that these waves can be arrested in space and analysed as for flowing streams with a steady wave-like surface deformation. Hitherto, this technique has been used almost exclusively to study progressive surface waves. For depths as great as 45 m, such wave-like deformations occur on the sea surface as a weakly coupled mirror-image of sandwaves on the seabed. By assuming that these surface deformations are real arrested progressive waves and not merely analogous, we now show that sand-dunes are related empirically by a small-amplitude surface wave dispersion equation. Furthermore, finite-amplitude wave theory (that is, steepest wave theory) is related to the steepest dunes on a stream bed where previously, for want of a better theory, steepest dune wavelengths have been explained in terms of fluid turbulence. In particular, if the dune wavelengths associated here with Gerstner's finite amplitude theory are combined in a manner similar to a beat effect, the predicted wavelengths are in good agreement with the lengths of marine sandwaves.

Allen¹ has shown that the deep-water approximation (Fig. 1, curve *a*) of the surface wave dispersion relation fits upper-regime antidune wavelengths well because of the strong in-phase steady surface waves associated with these bedforms. Furthermore, Kennedy² has noted that the steepness of antidunes is similar to that of deep-sea waves. This suggests that Gerstner's wave theory^{3,4} may be used to describe the form of antidunes. In Fig. 1, the dispersion equation (curve *b*), and curve *c* are the lower and upper limits, respectively, in Froude number *F*, to antidunes^{2,5}. *F* is defined by $F^2 = U^2/gh$ where *U* is the mean stream velocity, *h* is mean depth and *g* the acceleration due to gravity. Wavenumber *k* (Fig. 1) is defined by $k = 2\pi/L$, where *L* is a steady streamline wavelength, and in particular the wavelength of streamlines describing bedforms or surface water waves.

Previously, wave theory has been neglected as a means of explaining lower-regime bedforms (for example, dunes) because the associated steady surface wave amplitude *B* is invariably small. Instead, Yalin⁶ applied turbulence theory to explain steepest dune wavelengths. However, steady surface waves over

large dunes ($F < 0.2$) have been detected⁷ and this led us to believe that wave theory might also be applied to lower-regime bedforms.

Milne-Thomson³ has shown that small-amplitude wave theory may be applied to steady two-dimensional flow over a rigid sinuous bed of small amplitude *A* and wavenumber *k* so the surface wave amplitude *B* is given by

$$B/A = 1/(\cosh(kh) - (g/kU^2) \sinh(kh)) \quad (1)$$

The least depth *H* required for a water surface profile of wavenumber *k* to propagate at phase velocity *U* is

$$H = (1/k) \tanh^{-1}(khF^2) \quad (2)$$

At depth *H* the steady streamlines are straight. The shallow-water approximation ($L > 20h$) for equations (1) and (2) are

$$B/A = F^2/(F^2 - 1) \quad \text{and} \quad F^2 = H/h \quad (3)$$

B/A is positive or negative as U^2 is greater or less than C^2 where the phase velocity *C* is given by $kC^2/g = \tanh(kh)$. Equation (1) describes lower ($B/A < 0$) and upper ($B/A > 0$) regime bedforms. To test the hypothesis that lower-regime bedforms are also related to *F* by a generalized wave dispersion relation, we have fitted an expression of the form

$$kh(\beta F)^\alpha = \tanh(kh) \quad (4)$$

to the flume data of Guy *et al.*⁸ Figure 2 shows that equation (4) is supported by a linear trend giving a correlation coefficient significant at the 1% level for each sand size. Linear regression values of α and β were used to plot the curves shown in Fig. 1. Values of 1.2, 1.7 and 0.3 were calculated for α , β and r , respectively, using Stein's measurements from experiments⁹ with a sand size of 0.4 mm. The poor correlation coefficient *r* can be explained by data deficient in sand ripple measurements, as *F* was in the range 0.34–0.71. The dunes were generally not shorter than $2\pi h$ so the dune ratio $1/kh$ is virtually independent of *F*, since *F* tends to $1/\beta$ (Fig. 1, curves 1–3).

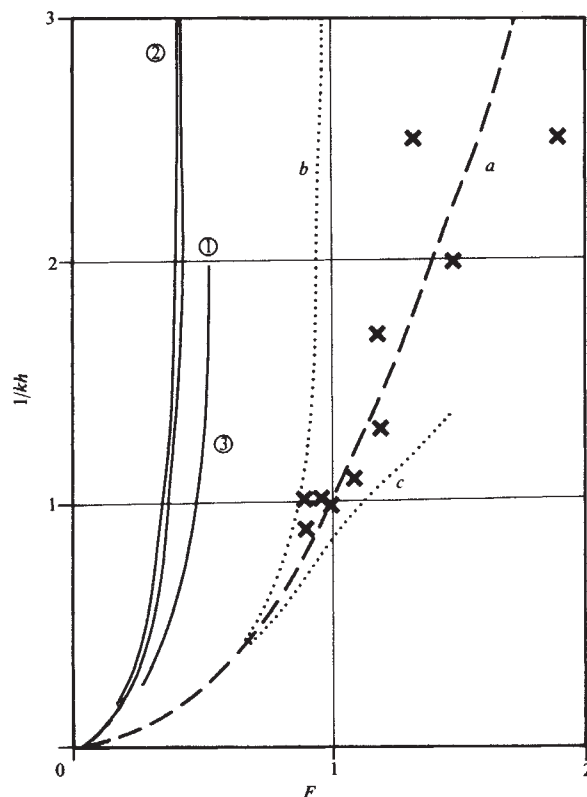


Fig. 1 The equations for *a*, *b* and *c* are, respectively, $khF^2 = 1$; $kC^2/g = \tanh(kh)$; and $khF^2 \tanh(kh) = 1$. The antidune measurements (x) are largely a sample from data collated by Kennedy². Curves 1, 2 and 3 are plots of equation (4) for respective sand sizes 0.19, 0.45 and 0.93 mm.

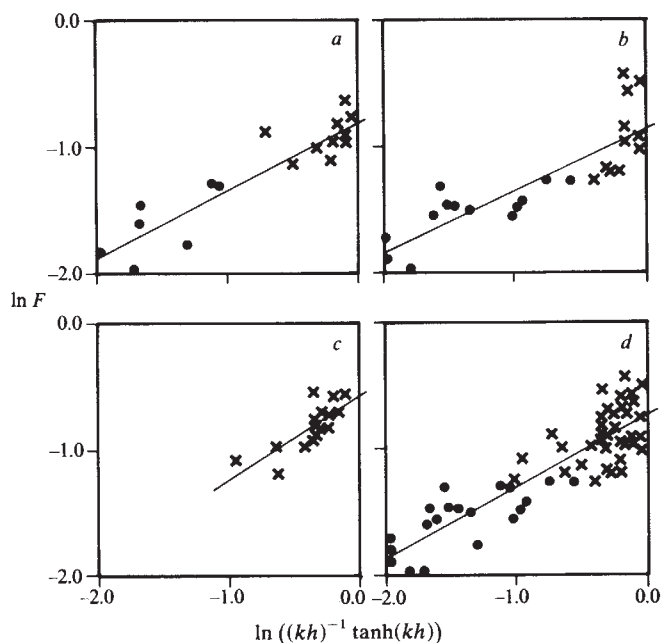


Fig. 2 Ripples (●) and dunes (×) as defined by Guy *et al.*⁸. The y intercept of the trend equals minus $\ln \beta$ and the slope equals $1/\alpha$. Numerical values for a – d are, respectively: 0.19 mm sand, $\alpha = 1.86$, $\beta = 2.23$, $r = 0.92$; 0.45 mm sand, $\alpha = 2.04$, $\beta = 2.36$, $r = 0.85$; 0.93 mm sand, $\alpha = 1.53$, $\beta = 1.80$, $r = 0.84$; the three grain sizes combined, $\alpha = 1.78$, $\beta = 2.09$, $r = 0.87$.

The above discussion has shown that small-amplitude wave theory may be applied to both upper and lower regime bedforms but not to the steepest bedforms. Yalin⁶ found from experimental evidence that the steepest dune wavelengths were given by

$$L \approx 2\pi h \quad (\text{that is, } kh \approx 1) \quad (5)$$

This result was explained in terms of macroscale fluid turbulence. However, the same result can be explained more simply in terms of finite-amplitude wave theory. To test relation (5), we have plotted the dune steepness against the dimensionless wavelength $1/kh$. Quadratic curves were fitted to Guy's data to give a single maxima only. Figure 3 shows that the maxima occur close to values of $1/kh = 1$ and the mean abscissa calculated for all the maxima is 0.99. Large fluvial dunes ($F < 0.2$) also have $kh \approx 1$ (Fig. 3d). Relation (5) was also tested using Stein's measurements. The regression analysis showed not a maximum but a minimum at the abscissa value of 2.13. Nevertheless, the relation is undoubtedly significant because 94% of the data, which started at $1/kh = 0.89$, lay with the negative slope of the curve which intersected $1/kh = 1$ at 0.042; that is, maximum dune steepness ($2A/L$) was typically 0.042.

Ideally, the transition from lower to upper regime bedforms is a plane bed because dune wavelength number $1/kh$ tends to infinity. Indeed, dune steepness does decrease (Fig. 3), as might be expected also by the critical depth H tending to h (that is, $F^2 \approx H/h \approx 1$). Minor disturbances to the fluid need distance to grow, and this explains why the intermediate regime is a steady plane bed in short flumes only¹⁰.

Most antidunes occur¹ when $F = 1$, which implies that steepest antidune wavelength ratios are $kh = 1$ (Fig. 1, curve a). Hence, steepest dunes and steepest antidunes have the same wavelength ratio $kh = 1$. Yalin's turbulence theory⁶ was applied to steepest dune wavelengths because bedforms were seen to grow in the lee of a discontinuity and the steady wave amplitude B was assumed by many to be negligibly small. However, typical large fluvial dune data are: $L = 85$ m, $2A = 1$ m, $U = 1$ m s⁻¹ and $h = 15$ m. Thus, from equation (1) it is possible to calculate an approximate surface wave height $2B$ of 6 mm. These small surface depressions, since they apparently interact with wind-waves, are detectable with satellite altimetry radar¹¹.

Table 1 Predicted values of L using equation (6)

General depth (m)		n	m	Predicted L	Observed mean L (m)
MHWS	MLWS				
14.5*	10.0	3	2	181	180
15.0†	5.0	3	1	94	70
25.0†	15.0	5	3	471	300
48.0‡	40.0	6	5	1,508	1,400

Data for the Bristol Channel entrance were based on a satellite radar-map; HMS Fawn, on passage to Bristol, confirmed the existence of the sandwaves, by echo-sounder.

* Start Bay, Devon¹⁵.

† Swansea Bay, Bristol Channel¹⁶.

‡ Bristol Channel entrance.

Gerstner's theory, by its simple geometry, relates easily with steepest bedform wavelengths and any appropriate local steepness can be generated. The pressure is constant along Gerstner's wavy streamlines; this is probably true for real stream flow except in the fluid space between depth H and mean depth h , where adverse pressure gradients promote flow reversal (Fig. 4).

Allen¹² has explained the sandwaves which form under tidal currents in shelf-seas by analogy with the ripple marks which form under wind-waves. Allen's explanation cannot account for the steady surface wave but it may account for sandwave shape. Free wave theory can be extended even further to explain why some sandwaves are longer than fluvial large dunes in the same general depth of water. If two large dune wavelengths are combined by a beat effect¹³ then the single least common multiple wavelength could explain the length of tidal sandwaves; thus,

$$L = m(2\pi h_1) = n(2\pi h_2) \quad (6)$$

where h_1 and h_2 are mean high water and low water spring tide depths and m , n are the smallest possible integers. Wavelengths predicted by equation (6) at four British continental shelf locations (Table 1) give good agreement with observation and a typical value of m or n is 3. Equation (6) predicts extremely long sandwaves (for example, $m/n \approx h_2/h_1 \approx 1$) in deeper water if an unusually large prime number satisfies the relation. However, probably because the sand distribution is limited, only the fluvial relation $kh \approx 1$ is possible in some shelf-seas; for example, some Alaskan sandwaves¹⁴ are ~250 m long in 65 m of water, where $m/n \approx 31/34$ would predict $L \approx 13$ km.

Several theories consider bedform amplitude growth as well as wavelength growth. We can speculate here that for lower-regime bedforms up to wavelength $L \approx 2\pi h$ steady surface waves are a consequence of bedform amplitude growth. The

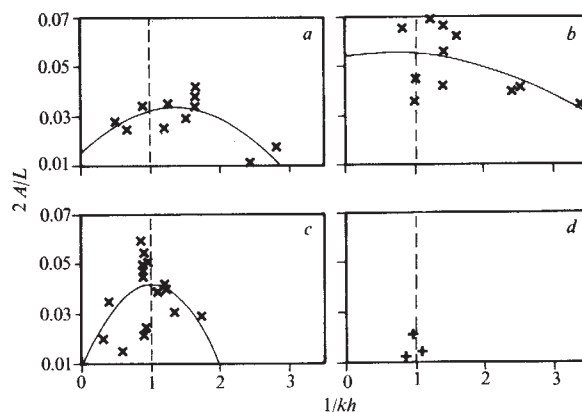


Fig. 3 Flume dune steepness (×) determined by Guy *et al.*⁸ for 0.19 (a), 0.45 (b) and 0.93 (c) mm sand. d, Fluvial large dune steepness (+) based on the measurements of Sundborg⁷, Smith and McLean¹⁷ and D. N. Langhorne (personal communication).

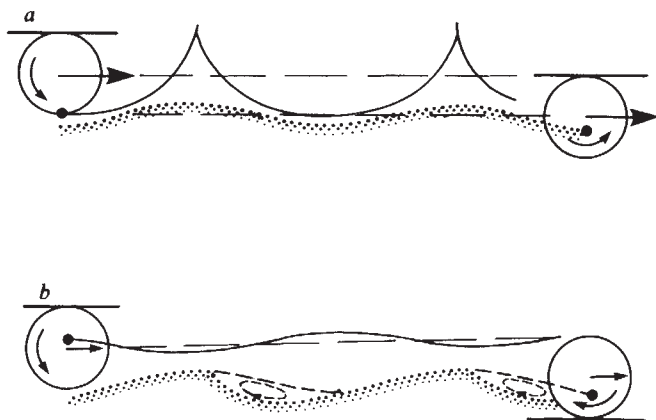


Fig. 4 A trochoidal wave^{3,4,13} is generated by a point of a disk rolled against a straight edge. A dimensionless wavelength $kh = 1$ implies that the radius of the disk equals h . The direction of rotation shown here, with constant translation, indicates the speed of the water particle. *a*, Although the steepness, $2B/L$, shown is almost double that of one particular laboratory¹⁰ antidune with sand size 0.45 mm, the height of the first crest that was about to break is the same. That is, $B = h = 120$ mm and the upstream surface trough was noted particularly to be lower than the downstream antidune crest. *b*, Flow separation is typical for steep dunes and sandwaves: the lee-eddy can be considered as part of the bedform². The speeds and heights of water particles would be more or less symmetrically distributed about the critical depth H .

steady streamlines down to depth H (equation (2)) are well suited to modelling by small-amplitude surface free waves. The datum at depth H can be regarded as a damping line to sand dune amplitude growth because B is always small.

Using this interpretation for the datum, dune amplitude (negative) growth is forced by unsteady out-of-phase surface waves. Once the datum is lower than h its role as a damping line is redundant: antidune growth begins (for example, see ref. 5) and the resonance predicted by equation (3) when $F \approx 1$ may be possible. Figure 4*a* shows resonance that may occasionally occur in laboratory flumes. Steep surface waves break in the upstream direction⁸ and not, as might be expected, in the downstream direction. This may be explained by fixing the frame of reference³ with the main stream so that the boundary profiles become progressive waves. The speed at each surface point probably depends on the precise depth to the datum and not the average depth H . Crests move faster than troughs and so break in the (upstream) direction.

The concept of constant motion relative to an inertial frame of reference fixed with the stream is, of course, not new, but we believe it has not been exploited intentionally in sedimentology. It has been shown to be effective for the analysis of the four most important bedform types—ripples, dunes, antidunes and sandwaves. Its full implication for sedimentology, however, needs further consideration. For instance, we have implied that its use provides an additional explanation, other than noting how sand grains are deposited, for the observation that the slightly steeper slopes of antidunes face upstream. Purely as a mirror-image in a datum, it follows that the steeper slopes of dunes face downstream, as indeed they do. These considerations are closely linked to the problem of selecting an unequivocal formula, from the many available formulae, for the mass transport rate of sand grains, given that important flow parameters, such as bed shear stress, are known.

Received 7 April; accepted 21 July 1981.

- Allen, J. R. L. *Physical Processes of Sedimentation* (Allen & Unwin, London, 1970).
- Kennedy, J. F. J. *Fluid Mech.* **16**, 521–544 (1963).
- Milne-Thomson, L. M. *Theoretical Hydrodynamics* (Macmillan, London, 1968).
- Lamb, H. *Hydrodynamics* (Cambridge University Press, 1919).
- Parker, G. J. *Hydraul. Div. Am. Soc. Civ. Engrs* **101**, (HY2) 211–221 (1975).
- Yalin, M. S. *Mechanics of Sediment Transport* (Pergamon, Oxford, 1977).
- Sundborg, A. *Geogr. Annl.* **38**, 126–316 (1956).

- Guy, H. P. et al. *Geol. Surv. Prof. Pap.* 462-I (U.S. Print Office, Washington DC, 1966).
- Stein, R. A. J. *Geophys. Res.* **70**, 1831–1842 (1965).
- Simons, D. B. et al. *Geol. Surv. Water-Supply Pap.* 1498A (U.S. Print Office, Washington DC, 1961).
- McLeish et al. *Mar. Geol.* (submitted).
- Allen, J. R. L. *Sedimentology* **27**, 317–324 (1980).
- Ippen, A. T. *Estuary and Coastline Hydrodynamics* (McGraw-Hill, New York, 1966).
- Bouma, A. H. et al. *Mar. Geotechnol.* **2**, 291–308 (1977).
- Langhorne, D. N. *Mar. Geol.* (in the press).
- Blackley, M. W. L. *IOS Rep.* **60**, (1978).
- Smith, J. & McLean, S. R. J. *Geophys. Res.* **82**, 1735–1746 (1977).

Radiocarbon dating and stratigraphic resolution in Welsh lateglacial chronology

Stephen Lowe

Geography Section, City of London Polytechnic, Calcutta House, Old Castle Street, London E1 7NT, UK
Department of Geography, Queen Mary College, Mile End Road, London E1 4NS, UK

Much attention has focused recently on the lateglacial period in north-west Europe, especially on the problems of climatic subdivision and chronostratigraphic definition^{1–3}. Correlation and classification of lateglacial stratigraphy relies heavily on radiocarbon dating, but because the lateglacial represents a relatively short interval of geological time (14,000–10,000 yr BP) with comparatively slow rates of sedimentation, the stratigraphic resolution achieved in radiocarbon dating may be questioned—a 5-cm thick sample may represent an age span significantly greater than the analytical uncertainty ($\pm 1\sigma$) associated with the radiocarbon measurement. Results are presented here for a radiocarbon-dated pollen stratigraphy based on sediments from Llyn Gwernan, a small freshwater lake situated on the northern flank of Calder Idris, 3 km south-west of Dolgellau, North Wales (grid ref. SH 703 159) where a remarkable thickness of lateglacial organic deposits preserves a detailed biostratigraphic record and permits a greater resolution of the lateglacial ¹⁴C time scale than is usual. The data are summarized in Fig. 1 and comparisons are made within the framework of previously published dates for Wales in Fig. 2.

Sediments have accumulated at the western edge of Llyn Gwernan, such that a gradual succession of plant communities across this infill have raised the bog level above that of general lake level. Organic lake muds were removed from the deepest part of this infill by piston corer and material from comparable horizons were bulked to obtain dates with low standard deviations on thin, precisely defined sediment slices (Fig. 1).

The date of $13,200 \pm 120$ yr BP (SRR-1705) marks the onset of organic sedimentation and provides a minimum age for late Devensian deglaciation of the site. This level in the profile is characterized by a marked upward trend in pollen of *Rumex*, with high percentage values for Gramineae and Cyperaceae. *Betula* values are low and *Juniperus* is absent. Only two comparable dated horizons have been published for Wales, of which a date of $11,660 \pm 140$ yr BP from Traeth Mawr, South Wales is thought to be anomalous⁴. At Gllanllynau, North Wales, dates of $12,556 \pm 230$ yr BP (seeds from basal organic sediments) and $14,468 \pm 300$ yr BP (moss-rich layer from bottom of basal minerogenic sediments) bracket the timing of thermal improvement based on fossil coleopteran succession⁵. Just as faunal, botanical and lithological responses have differing sensitivities to thermal improvement, the factors governing the thresholds for organic recruitment into different basins will be variable and depend on local climate, topography and geology. Direct comparison of pollen spectra from the basal organic horizons show *Juniperus* and *Rumex* rising almost simultaneously at Gllanllynau, whilst at Llyn Gwernan *Juniperus* first occurs some 30 cm higher in the profile than *Rumex*. Thus palynological evidence combined with the date of $13,200 \pm 120$ yr BP suggests the possibility of earlier organic sedimen-

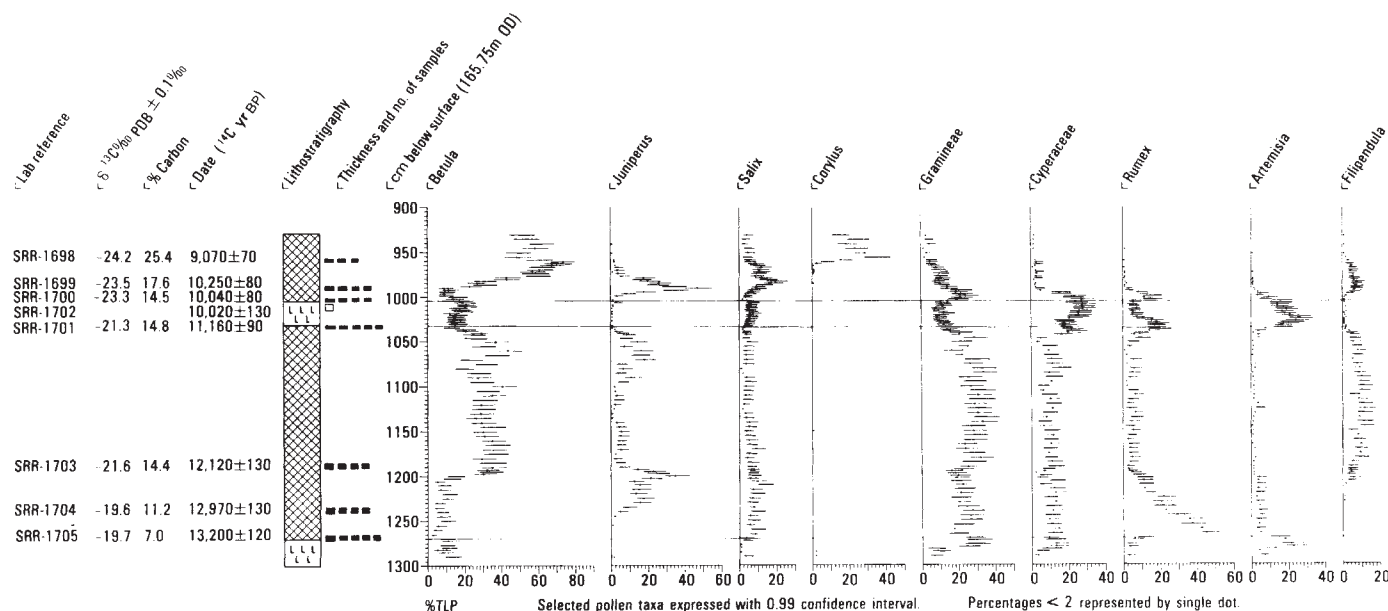


Fig. 1 Summary of ^{14}C data in relation to lithostratigraphy and pollen stratigraphy. Pollen sum per level = 300 total land pollen. $\square$, Organic lake mud; $\square$, clay.

tation at Llyn Gwernan than at Glanllynau, whilst still being consistent with much current evidence for timing of marked thermal improvement^{6,7}.

The first continuous pollen curve for *Juniperus* is dated at $12,970 \pm 130$ yr BP (SRR-1704). Although variable in occurrence and intensity, this relatively sudden expansion in *Juniperus* has been widely recognized throughout western Britain and explained as indicating response to improved climate⁸ or alternatively to processes of primary succession in an already temperate climate to a stage where it could flourish⁹. The few dates which exist for this phase indicate expansion in western Britain at $\sim 13,000$ – $12,500$ yr BP (ref. 8), but comparisons with many sites are hindered due to lack of precise definition of the expansion and the differing thicknesses of samples dated. Additional dates are required before the nature of this expansion can be examined.

The abrupt decline in the *Juniperus* curve following the expansion of *Betula* yielded a date of $12,120 \pm 130$ yr BP (SRR 1703). This decline has previously been interpreted as signifying the suppression of juniper scrub due to shading out by the superior competitive ability of birch¹⁰. However, the *Juniperus* decline is also a feature of many Irish sites, despite the relative absence of *Betula* and has been dated at $12,020 \pm 80$ yr BP (ref. 9). Furthermore, a growing body of evidence indicates that the increase in *Betula* pollen percentages is accompanied by the occurrence of coleoptera indicative of cooler summers⁷. Causal connection between climatic cooling and spread of birch woodland has been suggested¹¹ and seems to have been synchronous from southern to northern England, being dated at several sites to $\sim 12,200$ yr BP (ref. 7). A date of $11,300 \pm 300$ yr BP obtained for this event at Glanllynau may be erroneous due to possible contamination by fossil and modern roots¹². Although SRR 1703 falls close to the boundary of the Bolling/Older Dryas chronozone boundary¹, the palynological evidence does not support such an oscillation (Fig. 1). Likewise a date of $11,900 \pm 500$ yr BP for the decline in *Juniperus* at Nant Ffrancon which was interpreted as evidence for the Bolling oscillation¹³ could have alternative interpretations¹⁴. The radiocarbon date of $12,120 \pm 130$ yr BP from Llyn Gwernan may therefore mark the beginnings of climatic deterioration which eventually culminated in the truly severe cold period of the Loch Lomond (Younger Dryas) Stadial some 1,000 yr later.

The onset of the Loch Lomond (Younger Dryas) Stadial, as defined by an abrupt lithological change from organic lake muds to clay-rich minerogenic sediment, is dated at $11,160 \pm 90$ yr BP (SRR 1701) and corresponds with palynological changes from

tree, shrub and thermophilous herb pollen to a characteristic pollen assemblage of open habitat herbs, chiefly of Cyperaceae, *Rumex* and *Artemisia* (Fig. 1). This date is in close agreement with the age assigned to its onset in Scandinavia¹, Scotland² and the Lake District⁵ and cannot be used to support the hypothesis of later onset of stadial conditions in Wales⁴.

An age of $10,040 \pm 80$ yr BP (SRR 1700) was obtained for the lateglacial/Flandrian transition marking the end of the Loch Lomond Stadial and compares well with many similarly dated horizons from Wales, which in spite of varying standard deviations and different site characteristics, show uniformity in their mean values at $\sim 10,000$ yr BP (Fig. 2). However, a date for the peak in the *Juniperus* pollen curve, some 10 cm higher in the profile, gave an age of $10,250 \pm 80$ yr BP (SRR 1699) which viewed in the context of many other such dates in Britain (that is $\sim 10,500$ – $9,500$ yr BP (ref. 15)) could equally be considered reliable. A further date of $10,020 \pm 130$ yr BP (SRR 1702) was obtained on a sample of *Betula* wood recovered from the stadial clay some 2 cm lower in the profile than SRR 1700, which pollen evidence suggests was sampled *in situ*. We cannot differentiate between these dates at the $\pm 2\sigma$ level. Although problematic, the lateglacial/Flandrian boundary at Llyn Gwernan is dated at $\sim 10,000$ yr BP.

Expansion of *Corylus* at Llyn Gwernan is dated at $9,070 \pm 70$ yr BP (SRR 1698). This is in complete accord with evidence for the main expansion of *Corylus* at $\sim 9,000$ yr BP (ref. 15), but conflicts with the hypothesis of early invasion by *Corylus* from western Britain¹⁶, with the Cardigan Bay area possibly providing a refugium during the late Devensian¹⁷ (supported by a date of $9,747 \pm 220$ yr BP from Tregaron bog¹⁸). SRR 1698 compares

Table 1 Comparative data for Welsh lateglacial interstadial deposits

Site name (grid ref.)	Thickness of organic interstadial deposits (cm)	Age span of 1 cm of sediment (yr)	Age span represented by pollen count (yr)
Traeth Mawr ⁴ (SO 967257)	30	~ 67	~ 125
Glanllynau ¹² (SH 455373)	30	~ 67	~ 153
Nant Ffrancon ²⁴ (SH 632633)	140	~ 14	~ 71
Llyn Gwernan (SH 703159)	238	~ 8	~ 42

Ages have been derived assuming a constant sedimentation rate over a 2,000-yr period for the interstadial and are therefore approximate values drawn merely for comparison.

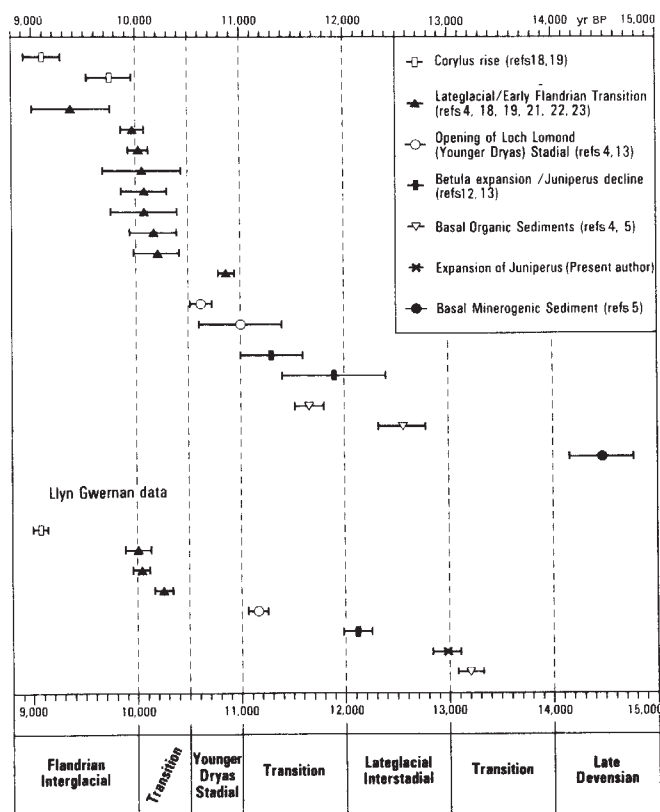


Fig. 2 Published lateglacial and early Flandrian radiocarbon dates with $\pm 1\sigma$ error bars from Wales. Stratigraphic terms according to Lowe and Gray³.

with similar dates from Nant Ffrancon ($9,098 \pm 180$ yr BP (ref. 19)), Coolteen, south-west Ireland ($9,055 \pm 95$ yr BP (ref. 9)) and Hawks Tor, Devon ($9,061 \pm 160$ yr BP (ref. 20)). Rather than providing evidence for early expansion in *Corylus*, collectively these dates suggest a remarkable degree of regional synchronicity for this phenomenon in south-west Britain.

In the comparisons drawn above, radiocarbon dates have been compared for assumed comparable events. Apart from general agreement in ages obtained for the lateglacial/early Flandrian transition $\sim 10,000$ yr BP, Figure 2 illustrates the paucity of lateglacial radiocarbon dates from Wales. Some of these dates lack precise definition and many have large standard errors. However, notwithstanding the inherent uncertainties associated with radiocarbon dating²⁵⁻²⁷ and the fact that comparable events need not have been synchronous, the most fundamental consideration affecting comparisons of ^{14}C dates concerns the relative resolution afforded by different sites. The unsatisfactory nature of the Welsh lateglacial chronology in terms of resolution and precise definition can in part be explained by differences in resolution between sites and the sampling techniques. Table 1 illustrates such variation in potential resolving power between Llyn Gwernan and the other available radiocarbon dated Welsh lateglacial profiles. Clearly the different characteristics of individual sites require different sampling procedures to obtain comparable precision in dating. Approximations of sedimentation rates and percentage carbon content of relevant horizons should therefore be evaluated before ^{14}C analyses to obtain an estimation of optimum sample thickness and mass for the required level of precision. The resolution afforded at Llyn Gwernan has therefore been an important factor in enabling the dating of sediments with negligible time spans which when bulked from comparable horizons have provided age determinations with low standard deviations and as such provide a significant contribution to the lateglacial chronology in Wales.

The advantages of sites such as Llyn Gwernan equally apply to the detailed studies of biostratigraphies and have wider impli-

cations for future lateglacial research throughout Britain and north-west Europe in general. Lateglacial profiles with good stratigraphic resolution should provide important information for testing the validity of major fluctuations in biostratigraphies. If such fluctuations are not apparent by detailed examination at 'high resolution sites', then the evidence for proposed fluctuations (with climatic implications) at sites of 'low resolution' may be questioned. Concentration of radiocarbon dating on sites containing good stratigraphic resolution is urged, as such sites will provide the best way of testing hypotheses concerning the synchronous or time-transgressive nature of lateglacial environmental changes.

I thank Drs J. M. Gray and J. J. Lowe for critically reading the manuscript, Dr D. D. Harkness for processing the radiocarbon samples, A. O. Lewis, R. M. Tipping and M. Ward for field assistance, and J. E. Cheney for aid with computer graphics. Financial support from the NERC and City of London Polytechnic is gratefully acknowledged.

Received 7 May; accepted 30 July 1981.

1. Mangerud, J., Andersen, S. T., Berglund, B. E. & Donner, J. J. *Boreas* 3, 109-126 (1974).
2. Gray, J. M. & Lowe, J. J. in *Studies in the Scottish Lateglacial Environment* (eds Gray, J. M. & Lowe, J. J.) 163-181 (Pergamon, Oxford, 1977).
3. Lowe, J. J. & Gray, J. M. in *Studies in the Lateglacial of North-West Europe* (eds Lowe, J. J., Gray, J. M. & Robinson, J. E.) 157-175 (Pergamon, Oxford, 1980).
4. Walker, M. J. C. *Nature* 287, 133-135 (1980).
5. Coope, G. R. & Brophy, J. A. *Boreas* 1, 97-142 (1972).
6. Ruddiman, W. F. & McIntyre, A. *Quat. Res.* 3, 117-130 (1973).
7. Coope, G. R. *Phil. Trans. R. Soc. B280*, 313-337 (1977).
8. Pennington, W. *Phil. Trans. R. Soc. B280*, 247-271 (1977).
9. Craig, A. J. *J. Ecol.* 66, 297-324 (1978).
10. Pennington, W. *Boreas* 4, 157-171 (1975).
11. Coope, G. R. & Joachim, M. J. in *Studies in the Lateglacial of North-West Europe* (eds Lowe, J. J., Gray, J. M. & Robinson, J. E.) 55-68 (Pergamon, Oxford, 1980).
12. Simpkins, K. *New Phytol.* 73, 605-618 (1974).
13. Burrows, C. J. *New Phytol.* 75, 167-171 (1975).
14. Moore, P. D. *Nature* 257, 356-357 (1975).
15. Smith, A. G. & Pilcher, J. R. *New Phytol.* 72, 903-914 (1973).
16. Deacon, J. *New Phytol.* 73, 1055-1063 (1974).
17. Moore, P. D. *New Phytol.* 71, 947-959 (1972).
18. Hibbert, F. A. & Switsur, V. R. *New Phytol.* 77, 793-807 (1976).
19. Switsur, V. R. & West, R. G. *Radiocarbon* 15, 156 (1973).
20. Brown, A. P. *Phil. Trans. R. Soc. B276*, 251-320 (1977).
21. Shotton, F. W. & Williams, R. E. G. *Radiocarbon* 15, 451-468 (1973).
22. Shotton, F. W., Williams, R. E. G. & Johnson, A. S. *Radiocarbon* 17, 255-275 (1975).
23. Handa, S. & Moore, P. D. *New Phytol.* 77, 205-225 (1976).
24. Seddon, B. *Phil. Trans. R. Soc. B244*, 459-481 (1962).
25. Shotton, F. W. *Nature* 240, 460-461 (1972).
26. Olsson, I. U. *Geol. Foren. Stockh. Forh.* 96, 311-320 (1974).
27. West, R. G. *Pleistocene Geology and Biology* (Longman, London, 1977).

Fragmentation of Asia in the Permian

M. W. McElhinny*, B. J. J. Embleton†, X. H. Ma‡ & Z. K. Zhang‡

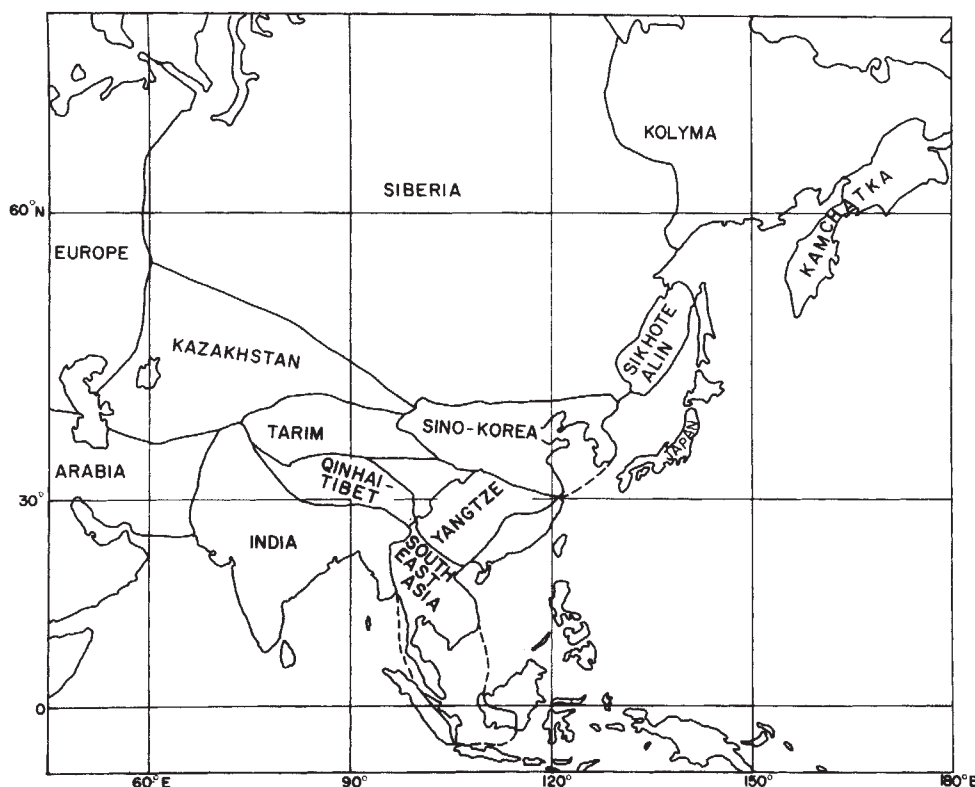
* Research School of Earth Sciences, Australia National University, Canberra, ACT 2600, Australia

† CSIRO Division of Mineral Physics, North Ryde, New South Wales 2113, Australia

‡ Chinese Academy of Geological Sciences, Beijing, Peoples Republic of China

Asia is a composite continent formed by the accretion of crustal blocks¹⁻³. A likely sequence of events is that Siberia collided with a Kazakhstan block which together then collided with Europe along the Urals in Permo-Triassic times². This was followed by the successive accretion of various blocks of south and east Asia to Siberia, culminating in the collision of India with Asia during the Cenozoic. Palaeomagnetic data have already demonstrated that the Kolyma and Sikhote Alin blocks of eastern Asia collided with the main continent along the late Mesozoic fold belts of Verkhoyansk and Sikhote Alin⁴. Furthermore, results from the Malay peninsula⁵ and Japan⁶ suggest that South-east Asia and Japan were situated near the Equator in Permian times and were therefore separated from the Asian continent. We report here some palaeomagnetic results from the Permian of the Sino-Korean and Yangtze blocks of China that show they were also situated near the Equator, were not in their present relationship to one another and were widely separated from Siberia.

Fig. 1 A possible subdivision of Asia into various palaeoblocks.



Surface geology suggests that Asia may be divided tectonically into blocks that once existed as separate plates^{3,7,8}. We have divided Asia into the late Palaeozoic blocks illustrated in Fig. 1. In the north the Soviet Union occupies the Siberian and Kazakhstan blocks with the Kolyma and the Sikhote Alin blocks to the east. The Cenozoic fold belts of Japan and Kamchatka lie on the eastern edge. A South-east Asian block comprises Indochina, the Malay peninsula, West Kalimantan (Borneo) and the north-east part of Sarawak. Palaeomagnetic results⁹ show that the Malay peninsula and West Kalimantan were a unit in Cretaceous times; we have tentatively extended this conclusion back to the late Palaeozoic.

The four palaeoblocks in China are named following Huang⁸. The Sino-Korean block is bounded by the Siberian block to the north and by the Qinling foldbelt to the south, separating it from the Yangtze block. To the west the Qinhai-Tibet block is bounded by the Himalayan foldbelt to the south and the Kunlun foldbelt to the north. In the north-west, the Tarim block is bounded by the Kunlun and Tien Shan foldbelts to the south and north respectively and the Nan Shan foldbelt to the east. It is not clear whether the Qaidam Basin to the south-east should be included as part of the Tarim block, nor whether the adjoining Ala Shan region to the east should be included as part of the Sino-Korean block as shown. In the extreme north-west of China, the Dzungar block is probably an eastern extension of the Kazakhstan block³ as indicated in Fig. 1. At present, the subdivision indicated in Fig. 1 is the simplest view; until detailed palaeomagnetic information is available, further subdivision is a matter of speculation.

Sampling of the late Palaeozoic (mainly Permian) type sequences was carried out at Ximing in the vicinity of Taiyuan, Shanxi Province (37.8°N, 112.3°E) on the Sino-Korean block and in the Emei Mountain region (29.6°N, 103.4°E) of the Yangtze block. Usually cores were drilled in the field using a portable drill, but some blocks were collected where the nature of the material made *in situ* drilling too difficult. Both the cores and the block samples were oriented using a Sun-compass orienting device.

In the Sino-Korean block, the sequence near Taiyuan is essentially flat-lying, ranging from Upper Carboniferous

through to the early part of the Lower Triassic. Although this whole sequence, covering ~1,300 m thickness, was sampled in the same detail (286 samples), we report here only from the redbed sequence in the Upper Permian. This sequence of red sandstones, siltstones and mudstones covers ~500 m thickness and takes in the uppermost member of the Upper Shihezi Formation (83 samples) and the overlying Shiqianfen Formation (43 samples). The type section has been subdivided into a sequence of numbered layers of varying thickness according to rock type. The results reported here cover layers 100–129 inclusive. The top of layer 129 corresponds to the top of the Shiqianfen Formation and is estimated to be at the top of the Upper Permian or very close to the base of the Lower Triassic.

In the Yangtze block the late Palaeozoic sequence near Emei Mountain is confined to the Permian and sedimentation extended into the early part of the Lower Triassic. The Permian sequence consists of Lower Permian limestones overlain in the Upper Permian by the Emeishan Basalt Formation and the Xuanwei Formation of redbeds. In the type section 17 flows of the basalt are exposed and have been folded so that the lowest 11 dip at 77° in a direction 64° (north-east) whereas the overlying six flows and the Xuanwei Formation dip at 87° towards a direction 62° (north-east). A second sequence was sampled ~50 km to the south-west. Here only seven flows of the basalt are exposed dipping at 46° in a direction 125° (south-east). The folding took place probably in the late Cretaceous and the structural combination sampled provided excellent prospects for the application of the fold test. As with the Sino-Korean sequence we report here only the results from the Upper Permian Emeishan Basalt and Xuanwei Formations. From the 24 flows 125 cores were drilled and 17 samples of the Xuanwei Formation were collected. Previously Wang *et al.*¹⁰ have reported results from 10 samples of Emeishan Basalt in the Emei Mountain region and from 8 samples of Permian sediments from Shandong in North China. In both cases the results were widely scattered (circle of confidence >25°) and, as no cleaning was carried out, it has been difficult to assess their significance.

Laboratory investigations were carried out at the palaeomagnetic laboratories at the ANU and CSIRO. All

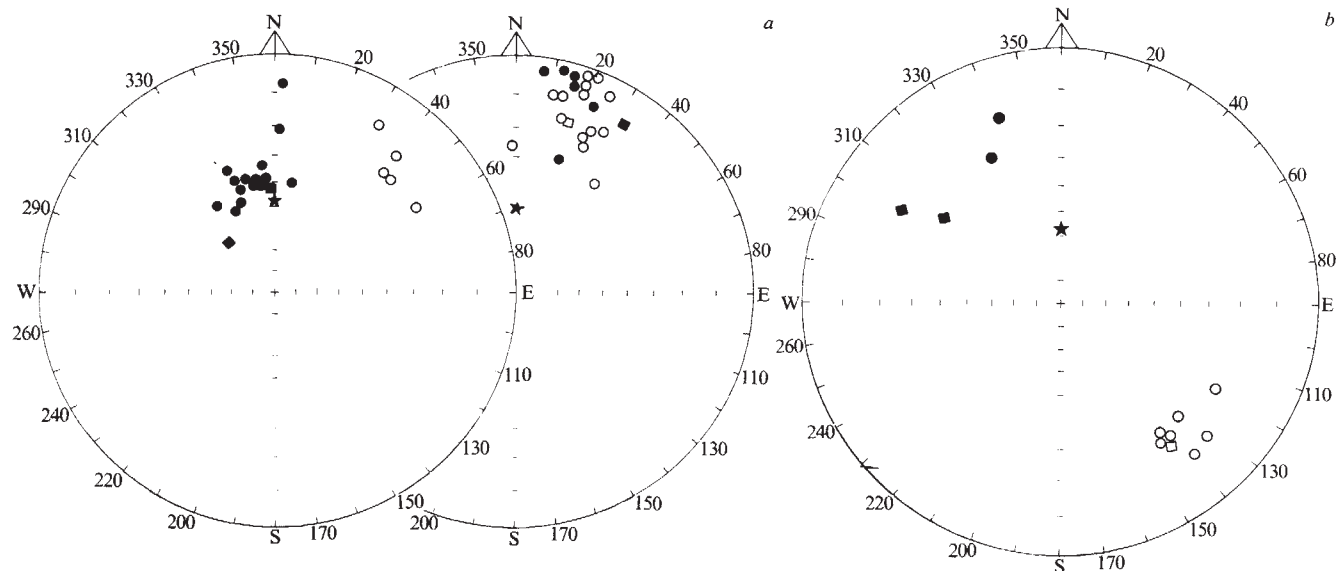


Fig. 2 Directions of magnetization. *a*, Observed in the Emeishan Basalt Formation (●○) and the overlying Xuanwei Formation redbeds (■□) before (left) and after (right) restoring the strata to the horizontal. *b*, Observed in the upper member of the Upper Shihezi Formation (●○) and overlying Shiqianfen Formation (■□). Stereographic projection, solid symbols on the lower hemisphere and open symbols on the upper hemisphere. ★, Direction of the present axial dipole field.

samples were subjected to stepwise thermal demagnetization, up to 670 °C in the case of the Xuanwei Formation and 570 °C in the case of the basalts. Some basalt samples were also treated in alternating magnetic fields. Thermal demagnetization seemed to be the most efficient process for cleaning. Data processing was carried out using orthogonal vector diagrams following Zijderveld¹¹ and these were analysed using the method of Kirschvink¹². The principal components deduced from this analysis generally revealed uniformly directed high temperature components of magnetization for each formation. Characteristic directions were observed in all but four basalt flows but in only six samples of the redbeds. In both sections a persistent NW–SE overprint magnetization was observed, clearly post-folding in age.

In the main type section the 14 *in situ* basalt directions lie very close to the direction of the axial dipole field direction (0°, +48.6°) or the present geomagnetic field direction (359°, +43°) before applying the bedding correction. However, this is not the case at the second section where the directions *in situ* diverge widely from the present field. On applying the bedding corrections the directions from the two sequences converge, confirming the successful application of the fold test (Fig. 2*a*). The mean *in situ* directions for the 14 flows in the type section is $D = 348.0^\circ$, $I = +39.7^\circ$, $k = 55.3$ and $\alpha_{95} = 5.4^\circ$. This mean direction is significantly different from that of the present geomagnetic and the axial dipole field, strongly suggesting they were not remagnetized in Recent times. The proximity to the present field direction must be regarded as coincidence.

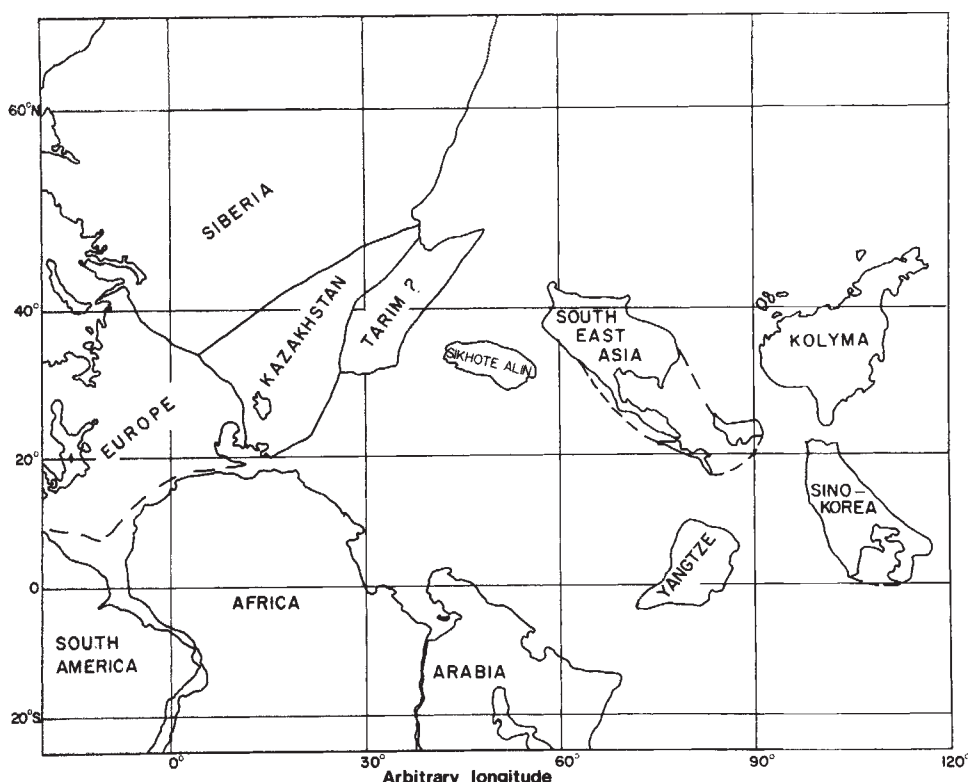


Fig. 3 Late Permian palaeogeographical map showing possible positions of the various blocks of Asia in the Tethys Ocean (or proto-Pacific). Relative longitudes are indeterminate so that the order in which the blocks are placed is arbitrary. The reconstruction of Gondwana with Laurasia is Pangaea B from ref. 16.

The significance of the fold test may be assessed in two ways. Following ref. 13, the ratio of precisions of the 22 site mean directions (Emeishan Basalt plus Xuanwei Formation) before ($k_1 = 5.5$) and after ($k_2 = 26.4$) applying bedding corrections are compared with F -ratio tables with $2(N-1)$ degrees of freedom. The ratio $k_2/k_1 = 4.8$ and this is very much higher than the value 2.0 for $F_{42,42}$ at the 1% level. This suggests that the fold test is successful at the 99% confidence level. A more sophisticated and rigorous test proposed by McFadden and Jones¹⁴ compares the lengths of the resultant vectors $\mathbf{R}_i$ from each of m limbs of a folded system with the overall resultant $\mathbf{R}$ to test whether or not they have a common mean direction. The statistic

$$P = \frac{(N-m) \sum \mathbf{R}_i - (\mathbf{R}^2 / \sum \mathbf{R}_i)}{(m-1) 2(N - \sum \mathbf{R}_i)}$$

is compared using F -ratio tables with $2(m-1)$ and $2(N-m)$ degrees of freedom. In the present case $N = 22$ and $m = 3$ as indicated in Table 1. Before folding, the statistic P has a value 21.2 which is well above the 5% significance point of 2.6 for $F_{4,38}$. The hypothesis of a common mean direction before folding may therefore be rejected. After folding, the statistic P reduces to 1.3 so that there is no reason to suppose that they do not have a common mean direction. The fold test is thus highly significant (Fig. 2).

Samples from the flat-lying redbed sequence near Taiyuan were subjected to successive thermal demagnetization to 670 °C. The Zijderveld diagrams showed that some rock types, particularly the sandstones, were strongly overprinted in the direction of the present geomagnetic field. The red mudstones, however, appeared to have little or no overprint components. After analysing the Zijderveld diagrams, characteristic components could only be determined for ~50% of samples. The characteristic directions revealed by principal component analysis have been averaged by layer and are shown in Fig. 2b. Overall means are given in Table 1.

In both the Taiyuan and Emei Mountain sequences the palaeomagnetic results show a basal single polarity zone in the Upper Permian overlain by a mixed zone of alternating polarities. We take this to indicate the top of the late Palaeozoic (Kiaman) Reversed Interval as found in rock sequences from around the world¹⁵. If this is correct then the basal directions in each sequence must be regarded as having reversed polarity. In terms of pole positions the Upper Permian of the Sino-Korean block has a mean (reversed) direction from 12 sites at $D = 138.4^\circ$, $I = -21.5^\circ$ and this corresponds to a North Geographic

Table 2 Permian palaeomagnetic poles and palaeolatitudes of the various blocks of Asia

Crustal block	Permian Pole	Central location	Palaeo-latitude	Ref.
Siberia/Kazakhstan	47°N 157°E	60°N 100°E	57.5°N	16
Kolyma	52°N 279°E	65°N 150°E	33.4°N	18
Sikhote Alin	18°N 196°E	48°E 137°E	33.9°N	19
Sino-Korea	44°N 358°E	38°N 115°E	9.8°N	This study
Yangtze	54°S 72°E	28°N 110°E	1.7°N	This study
Southeast Asia	57°N 152°E	10°N 105°E	30.8°N	5
Japan	—	35°N 136°E	5°S	6

Pole at 44.3°N, 357.8°E. Likewise the Upper Permian of the Yangtze block has a mean (reversed) direction from 22 sites at $D = 18.4^\circ$, $I = -6.5^\circ$, corresponding to a North Geographic Pole at 53.7°S, 72.1°E.

Both results indicate that the Sino-Korean and Yangtze blocks lie in low latitudes (11.1°N and 3.3°N respectively at the localities sampled) during the late Permian. This contrasts with the high latitudes ($> 50^\circ$ N) experienced by the Siberian block at this time^{7,16}. Thus both these major blocks of China were well separated from the Siberian block. Furthermore the Sino-Korean and Yangtze blocks were not in their present relationship to one another, and a relative rotation of 120° has occurred between them since the late Permian.

Table 2 summarizes the Permian palaeomagnetic poles that have been obtained from the various blocks of eastern Asia. Also indicated is the Permian palaeolatitude of a central location in each block. It is generally visualized that in a reconstruction of the Pangaea the northern part of Gondwana and the southern parts of Asia formed a wide wedge suggesting that a very wide ocean, the Tethys Ocean (or proto-Pacific), separated these regions. Palaeomagnetic data have already indicated that the six crustal blocks of Asia in Table 2 were widely separated from the Siberian block in the late Permian, and were located in low latitudes. Apparently, therefore, these crustal blocks must have occupied the centre of this large ocean.

Figure 3 shows a possible palaeogeographic reconstruction of Asia in late Permian times. We assume that the Siberia/Kazakhstan block has just collided with Europe and they have a common pole as shown in Table 2¹⁶. We assume also that Qinhai-Tibet either formed a part of Greater India, or else was situated in the Southern Hemisphere in the vicinity of India⁷. For simplicity Iran is included as part of Arabia. In reconstructing the relative position of Gondwana and Laurasia in late Palaeozoic times we have followed the scheme outlined by Irving¹⁶ in which there were two configurations of Pangaea, A and B. In Fig. 2b therefore the configuration is that of Pangaea B before its rearrangement into Pangaea A (similar to Wegener's reconstruction). The various blocks of Asia are then placed in their correct latitude and orientation but the relative longitudes are as yet unknown. Thus in Fig. 3 the relative order and spacing of the various blocks is purely arbitrary. Recalling that Permian results from Japan suggest equatorial latitudes as well as for this block, there was clearly a large amount of continental crust occupying the Tethys Ocean. Whether or not these blocks previously made up the supposed lost continent of Pacifica¹⁷ is not yet clear. However, if these blocks did make up Pacifica, they were not in the neighbourhood of Australia in the late Permian as has been proposed¹⁷, nor is it likely that they were associated with Gondwana.

This project is part of a cooperative program between the two palaeomagnetic laboratories in Australia and the Chinese Academy of Geological Sciences. It could not have been undertaken without the extensive assistance provided by the geological field teams in China. We thank Wang Zejiu, Vice-President of the Academy, and Team Leaders Wang Ber Lin, Shanxi Province and Tao Xien Lin, Sichuan Province for invaluable assistance.

Table 1 Mean directions of the characteristic magnetizations observed in Upper Permian rocks of the Yangtze and Sino-Korean blocks of China

Units combined	N	D	I	R	k	α_{95}
Yangtze Block (29.6°N, 103.4°E)						
Xuanwei Formation	2	24.0	-2.4	1.9323	14.8	—
Emeishan Basalt						
Uncorrected	20	4.1	+26.4	16.5031	5.4	15.5
Corrected for dip	20	17.8	-6.9	19.2866	26.6	6.4
Limb 1	7	15.7	-7.6	6.7649	25.5	12.2
Limb 2	9	16.0	-2.4	8.8955	76.6	5.9
Limb 3	6	25.3	-11.3	5.6415	14.0	18.5
Overall mean						
Uncorrected	22	2.3	+29.2	18.1816	5.5	14.6
Corrected for dip	22	18.4	-6.5	21.2032	26.4	6.2
Sino-Korean block (37.8°N, 112.3°E)						
Shiqianfen Formation	3	310.3	+24.2	2.9536	43.1	19.0
Upper Shihezi redbeds	9	141.0	-20.5	8.8113	42.4	8.0
Overall mean	12	138.4	-21.5	11.7274	40.3	6.9

Unit weight given to each defined sedimentary layer or lava flow. N , number of layers or flows; D , declination and I , inclination of mean direction; R , resultant of N unit vectors; k , precision parameter; α_{95} , circle of 95% confidence about the mean direction.

Received 30 March; accepted 24 June 1981.

1. Dewey, J. F. & Bird, J. M. *J. geophys. Res.* **75**, 2625–2647 (1970).
2. Hamilton, W. *Bull. geol. Soc. Am.* **81**, 2553–2576 (1970).
3. Burrett, C. F. *Earth planet. Sci. Lett.* **21**, 181–189 (1974).
4. McElhinny, M. W. in *The Western Pacific: Island Arcs, Marginal seas, Geochemistry* (ed. Coleman, P. J.) 407–414 (University of Western Australia Press, 1973).
5. McElhinny, M. W., Haile, N. S. & Crawford, A. R. *Nature* **252**, 641–645 (1974).
6. Hattori, I. & Hirooka, K. *Tectonophysics* **57**, 211–235 (1979).
7. Scotese, C. R., Bambach, R. K., Barton, C., van der Voo, R. & Ziegler, A. M. *J. Geol.* **87**, 217–277 (1979).
8. Huang, Chi-Ching (T. K.) *Eclogae Geol. Helv.* **71**, 611–635 (1978).

9. Haile, N. S., McElhinny, M. W. & McDougall, I. *J. geol. Soc. London*, **133**, 133–144 (1977).
10. Wang, T., Teng, H., Li, C. & Yeh, S. *Acta geophys. Sin.* **9**, 125–138 (1960).
11. Zijdeveld, J. D. A. in *Methods in Palaeomagnetism* (eds Collinson, D. W., Creer, K. M. & Runcorn, S. K.) 254–286 (Elsevier, Amsterdam, 1967).
12. Kirschvink, J. L. *Geophys. J. R. astr. Soc.* **62**, 699–718 (1980).
13. McElhinny, M. W. *Geophys. J. R. astr. Soc.* **8**, 338–340 (1964).
14. McFadden, P. L. & Jones, D. L. *Geophys. J. R. astr. Soc.* (in the press).
15. Irving, E. & Parry, L. G. *Geophys. J. R. astr. Soc.* **7**, 395–411 (1963).
16. Irving, E. *Nature* **270**, 304–309 (1977).
17. Nur, A. & Ben-Avraham, Z. *Nature* **270**, 41–43 (1977).
18. Pechersky, D. M. *Akad. Nauk SSSR Izv. Earth Phys. Ser.* 69–83 (1970).
19. Vlassov, A. Ya & Popova, A. V. in *Rock Magnetism and Palaeomagnetism* (ed. Vlassov, A. Ya) 333–339 (Siberian Academy of Sciences, Krasnoyarsk, 1963).

Is volcanic ash a pneumoconiosis risk?

Francis H. Y. Green, Val Vallyathan,
Michael S. Mentnech, James H. Tucker
& James A. Merchant

Division of Respiratory Disease Studies, Appalachian Laboratory for Occupational Safety and Health, National Institute for Occupational Safety and Health, Morgantown, West Virginia 26505, USA

Peter J. Kiessling, John A. Antonius
& Philip Parshley

Emanuel Hospital, 2801 N. Gantenbeim Avenue, Portland,
Oregon 97226, USA

In vitro studies of the mutagenic, fibrogenic and cytotoxic potential of Mount St Helens volcanic ash, and its likely effect(s) on immunological, antiviral and antibacterial defence mechanisms, have been reported elsewhere^{1,2}. These studies indicated that the ash has a moderate fibrogenic potential as it was markedly cytolytic in a sheep red blood cell haemolysis assay, had a stimulatory effect on W1-38 fetal lung fibroblasts and increased production of acid hydrolases by ash-exposed alveolar macrophages. We report here that in rats exposed intratracheally to volcanic ash, there was an acute pulmonary inflammatory response, followed by a granulomatous and fibrotic reaction which persisted until the end of the 6-month study period. We also examined the lungs of two heavily exposed humans and found lesions consistent with a reaction to volcanic ash. Although the ash does not seem to be highly toxic, these results suggest that precautions should be taken to minimize exposure of high-risk individuals.

Our *in vivo* experimental model used 10–12-week-old SPF Fischer 344 rats injected intratracheally with 10 mg volcanic ash in 0.25 ml sterile saline (0.9%). Control animals received the same volume of saline alone. The ash was a dry sedimented sample from the first volcanic eruption collected at Spokane (a city 340 km from Mount St Helens). Mineralogical analysis using X-ray powder diffraction (XRD) and scanning electron microscopy (SEM) combined with energy dispersive X-ray analysis (EDX) and light microscopy indicated that most of the particles belonged to the plagioclase class of minerals (sodium, calcium and aluminium silicates) and that the ash was similar in composition to sedimented samples analysed elsewhere³. Other minerals identified included albite, oligoclase, bytownite, labradorite, titanomagnetite, hornblende and silica. Total crystalline silica constituted 7.2% by weight, of which 4.2% was cristobalite and 3.0% was quartz. Stereological analysis showed the respirable fraction of the ash (<10 µm in diameter) to be 99% by count and 81% by weight.

Groups of 10 or less exposed or control rats were killed at days 1 and 7 of the study period and at 1, 3 and 6 months. At day 1, the lungs showed a focal acute inflammatory cell response at the level of the terminal bronchioles and alveolar ducts, the reaction corresponding to sites of maximum ash deposition. At 7 days, the reaction was still focal in nature but mononuclear cells now predominated. Animals examined at 1 and 3 months showed a widespread granulomatous reaction. The granulomas were composed of macrophages, lymphocytes and ash in a reticulin stroma; some showed a central necrotic area and contained foreign body type giant cells. At 6 months, the granulomas were larger and contained more collagen than observed earlier (Fig. 1a). SEM combined with backscattered electron imaging (BSI) and EDX confirmed that the mineral particles in the lung were similar in elemental composition to the original ash sample (Fig. 1b).

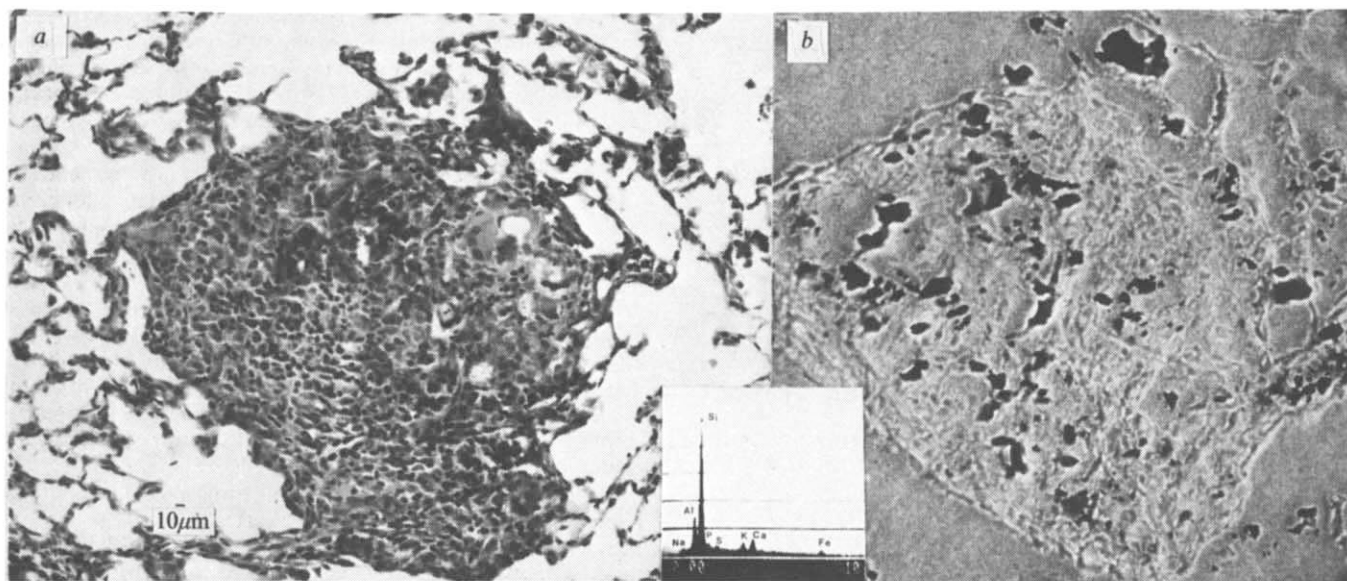


Fig. 1 a, Photomicrograph of rat lung examined 6 months after intratracheal injection of 10 mg volcanic ash. A granuloma composed of macrophages, lymphocytes, fibroblasts, giant cells and volcanic ash is seen, surrounded by relatively normal lung tissue (haematoxylin and eosin stain). b, Granuloma from adjacent section of the same animal examined in the BSI mode of the SEM. Volcanic ash particles in the granuloma appear black. Inset, EDX analysis of one of the particles detected by BSI, showing characteristic elemental peaks.

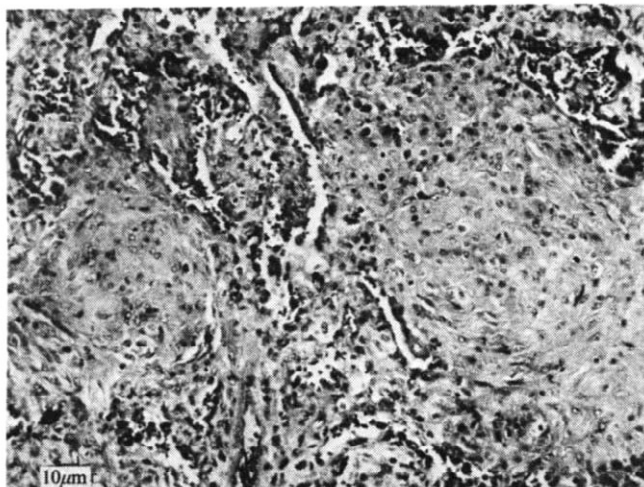


Fig. 2 Photomicrograph of lung from the logger who died 16 days after being rescued from Mount St Helens. Two organizing intra-alveolar granulomatous lesions are seen. Volcanic ash was identified in similar lesions by SEM combined with BSI and EDX (haematoxylin and eosin stain).

In addition, we have recently reviewed two unique cases of ash exposure in humans. The individuals were in a group of four loggers who were working in the area of the original blast on 18 May 1980. One died and the other three were admitted to the hospital with extensive burns; of these one survived with no apparent pulmonary dysfunction, and two, having generalized sepsis and pneumonia, died after 10–16 days⁴. In one of these two cases, the lungs showed intra-alveolar foci resembling the giant cell granulomatous lesions seen in the animal studies (Fig. 2), and the other showed an acute interstitial reaction with clusters of giant cells containing ash in the alveoli. Back-scattered scanning electron microscopic studies revealed inorganic particles in tissue sections which showed X-ray spectra similar to those of ash.

The laboratory studies reported here indicate that volcanic ash is moderately fibrogenic and should be considered a pneumoconiosis risk in heavily exposed individuals. Our preliminary observations of two humans heavily exposed to ash support this but certain qualifications must be made. First, the *in vivo* data are based on exposures far in excess of those likely to be encountered by the general population; second, these exposures were acute whereas future human exposures are likely to be to low concentrations of ash over relatively long periods. Finally, although the lesions seen in the loggers' lungs were similar in some respects to those in the animals, confounding factors such as burn injury, therapy and infection made it difficult to interpret the pathogenesis of the lesions.

The granulomas produced in experimental animals resembled those associated with mixed mineral dust exposure⁵, which are considered to represent a reaction to the crystalline forms of silica modified by other silicate minerals or oxides of iron⁵. Although non-quartz silicates, by themselves, are considered to be only weakly fibrogenic, a previously unrecognized pneumoconiosis was recently reported in California farm workers in which silicate minerals were implicated in the pathogenesis of pulmonary fibrosis⁶. Further work is required to determine which ash component(s) exert the fibrogenic effect. Based on these studies, and in view of uncertainty about future volcanic activity, we think precautions should be taken to monitor and minimize ash exposure in those likely to be heavily exposed.

Inherited retinal dystrophy in RCS rats: a deficiency in vitamin A esterification in pigment epithelium

E. R. Berman*, N. Segal*, S. Photiou*, H. Rothman* & L. Feeney-Burns†

* Ophthalmic Biochemistry Unit, Hadassah University Hospital, Jerusalem, Israel

† Department of Ophthalmology, University of Missouri School of Medicine, Columbia, Missouri 65201, USA

The RCS (Royal College of Surgeons) mutant strain of rats¹ has often served as a model for the study of inherited retinal degenerations. Lamellar outer segment debris accumulates early in the course of the dystrophy, probably as a consequence of a markedly diminished ability of the retinal pigment epithelium (RPE) to phagocytose photoreceptor disks^{2–4}. The RPE seems to be the primary site of the genetic defect⁵, and faulty recognition mechanisms in the RPE cell membranes may be involved in the disease process. The observation of an unusually active retinol esterifying enzyme in the microsomal fraction of cattle RPE⁶, that may form part of a vitamin A cycle in this cell layer⁷, prompted us to investigate retinol metabolism in the RPE of the RCS rat. We have now found that in dystrophic animals, a pronounced defect of retinol esterifying enzyme is apparent as early as the 10th postnatal day, and is accompanied by a small excess of retinol in the RPE.

Dystrophic, pink-eyed, tan-hooded rats of the RCS strain and control rats, either albinos of the Wistar strain or the RCS-rdy⁺ congenic strain, were reared from birth in 12-h cyclic illumination of less than 2 ft-candles. The animals were killed 3 h or more after the onset of fluorescent laboratory light (20–30 ft-candle). All RPE homogenates were prepared from age-matched litters of 5–15 animals by enucleating the eyes and, under a dissecting microscope, removing the anterior portions and the neural retina. The posterior eye cups were sonicated (Braunsonic 300S, microprobe) for 15 s in 0.9% NaCl and then discarded. This homogenate was used for all experiments. Sonication effectively removes all the RPE cell layer plus the adjacent choroidal tissue in both RCS and control rats (Fig. 1). The sclera remains essentially unaffected. The 10-day old RCS eye cup is morphologically similar to the control, whereas the 20-day old RCS preparations contain strongly adhering debris which cannot be surgically removed in animals of this age.

We have chosen only dim cyclic illumination for these studies so as to avoid any harmful effects of prolonged exposure to high intensity light, to which albino rats seem especially sensitive⁸. Previous measurements of retinol in the 'pigment layers' of fully light-adapted RCS rats showed an increase of 50–100% over that present in albino controls⁹. We have re-examined this finding under our lighting conditions because the absolute level of vitamin A in the pigment epithelium will vary according to the degree of light adaptation¹⁰. Retinol and retinyl ester were measured in the RPE of 10–12-day-old rats but not in older animals because of artefacts that would be introduced owing to the lamellar debris present (Fig. 1e). Table 1 shows an increased retinol content of 25–30% in dystrophic rats and a concomitant decrease in retinyl ester. The raised level of retinol found in the RCS pigment epithelium under our lighting conditions is less than the excess reported by Reading⁹ in animals of comparable age, and when subjected to statistical analysis (Student's paired *t*-test), the differences were not significant ($P > 0.05$).

The relatively small excess of retinol in RCS pigment epithelium probably does not arise from enhanced lability of photoreceptor rhodopsin⁹, because the bleaching properties of the visual pigment seem to be normal^{2,11}. Retinol, however, also

Received 16 June; accepted 20 July 1981.

- Green, F. H. Y. *et al.* in *Inhaled Particles* Vol. 5 (ed. Walton, W. H.) (Pergamon, Oxford, in the press).
- Olenchok, S. A. & Mull, J. C. *Life Sci.* **28**, 791–794 (1980).
- Fruchter, J. S. *et al.* *Science* **209**, 1116–1125 (1980).
- Parchman, F. *Outreach* **1**(2), 1–32 (1980).
- Parkes, W. R. in *Occupational Lung Disorders*, 194–197 (Butterworths, London, 1974).
- Sherwin, R. P., Barman, M. L. & Abraham, J. L. *Lab. Invest.* **40**, 576–582 (1979).

reaches the pigment epithelium from the circulation¹², and this is undoubtedly an important source of vitamin A in this cell layer, especially in 10–12-day-old rats whose eyes are still closed and whose visual cycle is not yet fully operative².

The possibility of a defect in retinol esterifying enzyme was tested by examining this activity in four age-matched groups of control and RCS rats. The results (Fig. 2) show a striking deficiency in all groups examined. No esterification was detectable in 10-day-old RCS rats except at low substrate concentrations, whereas at a comparable age, measurable activity (~ 0.6 nmol of ester formed per h per mg protein) was always observed in the controls; 12-day-old RCS rats showed only about 15% of normal activity. In most age groups, at substrate concentrations approaching or reaching saturation for the

control enzyme, that is, $> 1.5\text{--}2\text{ }\mu\text{M}$, the RCS enzyme exhibits only about 15–20% of normal activity. Double reciprocal plots of the velocity curves in Fig. 2 corroborated this, showing that the apparent V_{max} was about four to five times higher in the controls than in the RCS rats. On the other hand, the apparent K_m values for retinol esterification were similar in the RCS and control rats, both being in the range of $0.5\text{--}1.5 \times 10^{-6}\text{ mol l}^{-1}$.

The 20–30-day-old RCS pigment epithelium had relatively higher esterifying activity than that found in the other age groups of RCS rats, approaching $\sim 35\text{--}45\%$ of controls, especially at low substrate concentrations (Fig. 2). It was suspected that some of this activity could have derived from the adhering lamellar debris (Fig. 1e), which begins to accumulate early in the course of the retinal dystrophy^{2–4} and reaches its maximum

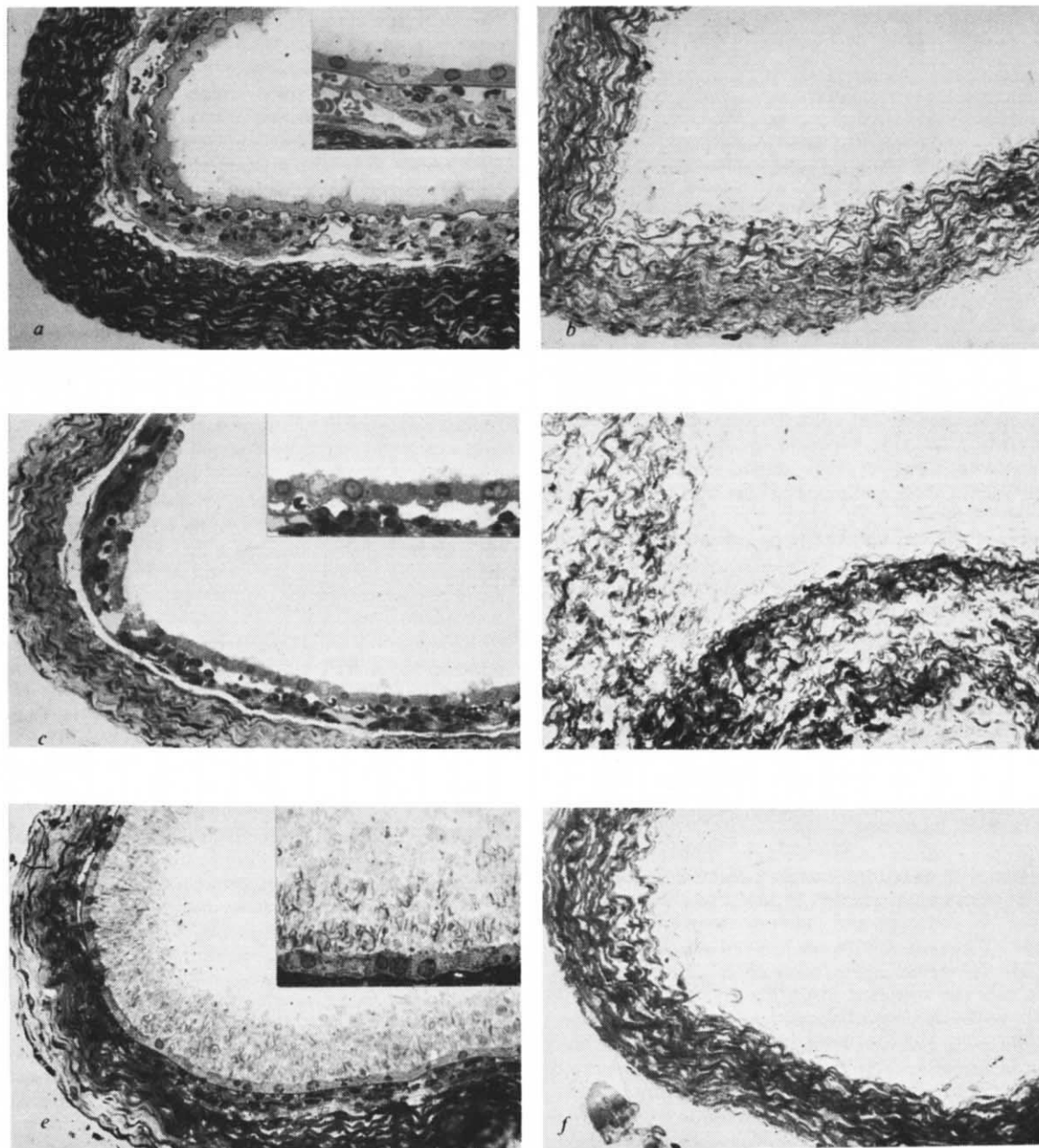


Fig. 1 Semi-thin ($1.0\text{ }\mu\text{m}$) sections of posterior eye cups of 10-day-old and 20-day-old RCS and control rats after removal of neural retina. *a*, Eye cup from 10-day-old control before sonication ($\times 255$). Inset: appearance of RPE and choroid at higher magnification ($\times 360$). *b*, Control eye after sonication, showing removal of RPE and choroid, with sclera essentially intact ($\times 255$). *c*, 10-day-old RCS eye cup before sonication. Note that its morphology is similar to that of the control eye cup ($\times 255$). Inset: 10-day-old RCS eye cup at higher magnification ($\times 360$). *d*, 10-day-old RCS eye cup after sonication is similar in appearance to the control. *e*, 20-day-old RCS eye cup before sonication. Note the layer of outer segment debris adhering to the RPE ($\times 255$). Inset: appearance of RCS posterior eye cup at higher magnification ($\times 360$). *f*, 20-day-old RCS eye cup after sonication, showing removal of RPE, choroid and lamellar debris, but leaving the sclera essentially intact ($\times 360$).

between about the 25th and 40th postnatal days^{2,13}. To test this possibility, the photoreceptor debris was isolated from 20–30-day-old RCS rat retinas by the method of Delmelle and co-workers¹⁴ and, using the assay described in Fig. 2 legend, it was found to have retinol esterifying activity of 1.6–3.7 nmol per h per mg protein. The amount of lamellar debris in pigment epithelial homogenates can be estimated from rhodopsin content¹⁴ and these measurements showed that in 20–30-day-old RCS rats, ~8–25% of the total photoreceptor debris adheres to the pigment epithelium, the absolute amount varying with the dissection technique. A simple calculation shows that the debris present in RPE homogenates of RCS rats could contribute ~20–50% of the total retinol esterifying activity detected. Even though the activity in debris and RPE need not be due to the same enzyme, the precise extent of esterifying deficiency in the RPE cannot be accurately assessed in the presence of lamellar debris.

The interphotoreceptor debris in the RCS rat begins to disappear from about the 40th postnatal day, and is virtually undetectable by the 96th day¹³. However, the pigment epithelial cell layer remains essentially intact during most of this period, notwithstanding certain areas of deterioration that can be observed by light microscopic examination¹³. In spite of the technical problems caused by accumulated debris and a slowly deteriorating pigment epithelium, it was nevertheless important to examine this activity in the RPE cell layer of completely mature RCS rats so as to exclude the possibility of delayed development of the retinol esterifying enzyme. We have found that in 52–55-day-old RCS rats, the debris that remains is less firmly adherent to the pigment epithelium and by careful dissection can be completely removed. Moreover, as shown in Fig. 3, the RPE appears normal morphologically at this age. It is apparent from the kinetic data in Fig. 2 that the enzyme does not, in fact, 'develop late'; rather, the same degree of deficiency of retinol esterification is present in mature RCS rats as in young

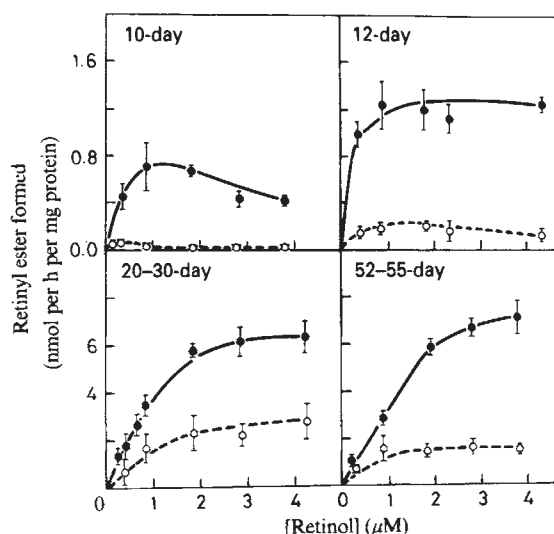


Fig. 2 Esterification of retinol as a function of substrate concentration in homogenates of pigment epithelium from dystrophic RCS rats (○) and controls (●). Homogenates were obtained as described in the text, the yield of protein being 150–300 μg per rat eye, depending on the age of the animal. All operations were carried out in dim light. The assay mixture of 400 μl contained 50–150 μl of freshly prepared pigment epithelial homogenates (125–150 μg of protein), 50 μmol of citrate-phosphate buffer pH 7.5, and ³H-retinol (2.21 Ci mmol⁻¹, NEN) in 10–20 μl ethanol at concentrations of 0.375–4.4 μM. Unlabelled retinol (prepared as described in Table 1 legend), was added as carrier so that the amount of radioactivity in each sample incubated was in the range 25,000–50,000 c.p.m. After 0.5 h incubation at 30 °C, the retinyl ester and retinol were extracted as described in Table 1 legend and elsewhere^{6,21}. The fatty acid moieties of the esters formed during incubation were not identified, but it may be assumed that they are either palmitate, stearate, or both²². The 3% and 50% diethyl ether/light petroleum effluents from the alumina columns were collected directly into counting vials, evaporated under nitrogen and counted in a toluene-based fluor in a Packard Tri-Carb liquid scintillation counter. A control consisting of buffer and substrate, without enzyme, was run with each experiment. The small amount of radioactivity in the 3% diethyl ether/light petroleum fraction (that is, retinyl ester) was subtracted from the total ester formed during incubation of labelled retinol with tissue samples. Esterifying activity was calculated as nmol retinyl ester formed per h per mg protein. All points represent the means ± s.e.m., *n* ≥ 3.

Table 1 Retinol and retinyl ester in the pigment epithelium of 10–12-day-old RCS and control rats

Strain	Retinol	Retinyl ester
	ng per eye	
RCS	0.42 ± 0.04	6.03 ± 0.96
Control	0.33 ± 0.06	7.76 ± 0.85

Pigment epithelial cell homogenates were obtained by sonication of eye cups after removal of the neural retina, using 15–18 animals for each experiment. All rats were kept in the lighting conditions described in the text. The retinol and retinyl ester were extracted in dim light by addition, with mixing, of an equal volume of ethanol, followed by extraction with 5 vol. of light petroleum (b.p. 35–60 °C). After brief centrifugation, the light petroleum was removed and the water-ethanol phase extracted twice more. The combined light petroleum extracts were concentrated to a small volume and chromatographed on 0.7 × 3 cm columns of 6% water-weakened alumina. The column was first washed with light petroleum, and 3% and 50% diethyl ether in light petroleum were applied to elute retinyl ester and retinol, respectively. Analytical grade or redistilled solvents were used throughout. The effluents were concentrated to small volumes, but not to dryness. The final volumes were equalized to 1 ml with light petroleum, and the concentrations of retinol and retinyl ester determined fluorometrically (325 nm excitation and 470 nm emission) in an Aminco SPF-125 spectrofluorometer equipped with a xenon lamp and integrated timer and printer. The instrument was standardized with quinine sulphate (0.2 μg ml⁻¹) in 0.05 M H₂SO₄ and the samples were read against solvents carried through the same procedure as the samples. All-*trans*-retinol, freshly prepared in the dark by reduction of all-*trans*-retinal (Sigma) with sodium borohydride, was used as the standard. Approximately 10 mg of sodium borohydride were added to a solution of 5 mg ml⁻¹ of retinal in ethanol. After standing for 3–5 min, 2 ml of water and 5 ml of light petroleum were added successively. The mixture was vortexed and centrifuged, and after twice washing the water-ethanol phase, the combined light petroleum extracts were concentrated under a stream of nitrogen and stored in nitrogen in a tightly fitting glass-stoppered tube. The minimum detectable amount of vitamin A was 8 ng. The data shown are the means ± s.e.m., *n* = 3.

10–12-day-old animals. A specific enzymatic defect in the pigment epithelial cell layer that is demonstrable from day 10 (when this cell layer is morphologically and functionally normal) until adulthood supports the view that a deficiency of retinol esterifying enzyme in the pigment epithelium may be the primary inherited defect in the RCS rat. The finding that retinyl ester is lower, but not totally absent, in the pigment epithelium of the RCS rat does not contradict this view because the major defect is in the rate at which the ester is formed at high substrate concentrations. At lower substrate concentrations (Fig. 2) the activity in the dystrophic rat may be sufficient to produce retinyl ester in amounts similar to those found in controls.

That the esterifying deficiency was not due to an inhibitory substance in the RCS homogenates was evident from 'mixing' experiments. Adding graded amounts of RPE from RCS rats to homogenates of controls caused no inhibition; instead, small increments in esterification corresponding to the individual activities contributed by both sources were found. Other aspects of RPE metabolism in the RCS rats are normal; for example, the retinol-binding protein in pigment epithelial cytosol of the RCS rat is indistinguishable from that in the controls¹⁵, and the receptors for serum retinol-binding protein also show no abnormalities in the RCS rat¹⁶. Essner and co-workers¹⁷ have shown that the lysosomes in RCS pigment epithelium are structurally intact and their enzyme content, as measured cytochemically, is normal. We have confirmed this by direct biochemical assay¹⁸ of *N*-acetylglucosaminidase, 4-methylumbelliferyl palmitate hydrolase, acid phosphatase, β-galactosidase and α-mannosidase in our RPE homogenates.

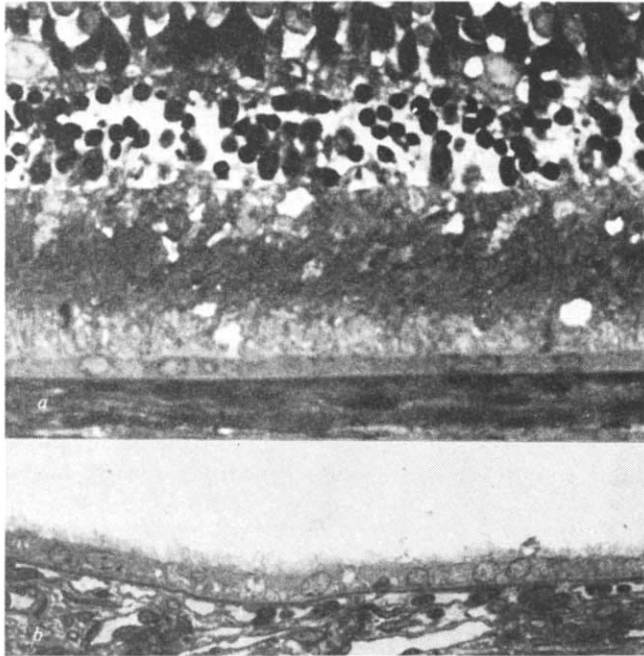


Fig. 3 Semi-thin ($1.0\ \mu\text{m}$) sections of a 52-day-old RCS rat retina. *a*, Before removal of neural retina. Note that the photoreceptors have pycnotic nuclei. Outer segment debris is still present, but RPE morphology appears normal. *b*, After removal of neural retina, little or no lamellar debris remains adherent to the RPE. These light micrographs are representative of most of the areas examined in serial sections of optic cups from RCS rats of this age. $\times 620$.

A metabolic block in vitamin A metabolism in the retinal pigment epithelium is one of several hypotheses proposed by LaVail to explain the photoreceptor cell death in the RCS rat¹⁹. In view of the known membranolytic properties of retinol and some of its derivatives (excluding retinyl ester)²⁰, it is tempting to speculate that a deficiency of esterifying enzyme and a concomitant increase in retinol, albeit a small one, could have a role in the pathogenesis of the retinal dystrophy in the RCS rat. This need not, however, be through direct action on lysosomal membranes of the RPE⁹, but rather on other membrane systems in the neural retina or RPE, especially during their early critical period of development.

We thank Drs Matthew LaVail and Michael Hall for breeding pairs of RCS and RCS-*rdy*⁺ rats, and Dr Marsha Kaitz for supplying us with animals from her colony. This work was supported by the US-Israel Binational Science Foundation, Jerusalem, the Eye Research Foundation of Missouri, the NIH and the Chief Scientists' Office, Ministry of Health, Israel.

Received 9 March; accepted 3 July 1981.

- Bourne, M. D., Campbell, D. A. & Tansley, K. *Br. J. Ophthalmol.* **22**, 613–623 (1938).
- Dowling, J. E. & Sidman, R. L. *J. Cell Biol.* **14**, 73–109 (1962).
- Herron, W. L. Jr, Riegel, B. W., Myers, O. E. & Rubin, M. L. *Invest. Ophthalmol.* **8**, 595–604 (1969).
- Bok, D. & Hall, M. O. *J. Cell Biol.* **49**, 664–682 (1971).
- Mullen, R. J. & LaVail, M. M. *Science* **192**, 799–801 (1976).
- Berman, E. R., Horowitz, J., Segal, N., Fisher, S. & Feeney-Burns, L. *Biochim. biophys. Acta* **630**, 36–46 (1980).
- Berman, E. R., Segal, N., Schneider, A. & Feeney, L. *Neurochemistry* **1**, 113–122 (1980).
- Kuwabara, T. in *The Retinal Pigment Epithelium* (eds Zinn, K. M. & Marmor, M. F.) 293–313 (Harvard University Press, 1979).
- Reading, H. W. *J. med. Genet.* **7**, 277–284 (1970).
- Dowling, J. E. *Nature* **188**, 114–118 (1960).
- Chaitin, M. H. & Williams, T. P. *Expl Eye Res.* **24**, 553–558 (1977).
- Robison, W. G. Jr & Kuwabara, T. *Invest. Ophthalmol. vis. Sci.* **16**, 1110–1117 (1977).
- LaVail, M. M. & Battelle, B. A. *Expl Eye Res.* **21**, 167–192 (1975).
- Delmelle, M., Noell, W. K. & Organisciak, D. T. *Expl Eye Res.* **21**, 369–380 (1975).
- Wiggert, B. O., Bergsma, D. R. & Chader, G. J. *Expl Eye Res.* **22**, 411–418 (1976).
- Bok, D. & Heller, J. *Invest. Ophthalmol. vis. Sci.* **19**, 1405–1414 (1980).
- Essner, E., Gorin, G. M. & Griewski, R. A. *Invest. Ophthalmol. vis. Sci.* **17**, 278–288 (1978).
- Rothman, H., Feeney, L. & Berman, E. R. *Expl Eye Res.* **22**, 519–532 (1976).
- LaVail, M. M. in *The Retinal Pigment Epithelium* (eds Zinn, K. M. & Marmor, M. F.) 357–380 (Harvard University Press, 1979).
- Goodall, A. H., Fisher, D. & Lucy, J. A. *Biochim. biophys. Acta* **595**, 9–14 (1980).
- Berman, E. R., Segal, N. & Feeney, L. *Biochim. biophys. Acta* **572**, 167–177 (1979).
- Alvarez, R. A., Bridges, C. D. B. & Fong, S.-L. *Invest. Ophthalmol. vis. Sci.* **20**, 304–313 (1981).

Retinal bipolar cells with double colour-opponent receptive fields

A. Kaneko & M. Tachibana

Department of Information Physiology, National Institute for Physiological Sciences, Myodaiji, Okazaki 444, Japan

Cells with colour opponency in each of two spatially antagonistic parts of their receptive fields, have been found at various stages of the visual pathways in several animal species^{1–8}. These cells, which are said to have 'double colour-opponent receptive fields', are thought to play a major role in the analysis of coloured images by enhancing simultaneous colour contrast². This stage of information processing was heretofore considered to begin at the level of retinal ganglion cells^{1–5}. Here we report that, in the carp at least, this stage occurs earlier, as we have found double colour-opponent receptive fields among the retinal bipolar cells of this animal.

Using techniques described previously⁹, intracellularly recorded responses of bipolar cells were measured in retinas isolated from carp (*Cyprinus carpio*), superfused with an artificial saline and partially light-adapted by a constant background ($2\ \text{cd m}^{-2}$). Eighty-five cells were studied in a series of experiments designed to analyse more fully the receptive field structure of bipolar cells (A.K. and M.T., in preparation). Many of the cells studied were identified from their response characteristics, that is, sustained and graded voltage changes, lack of action potentials and antagonistic centre surround organization of their receptive fields. These identifications were confirmed by

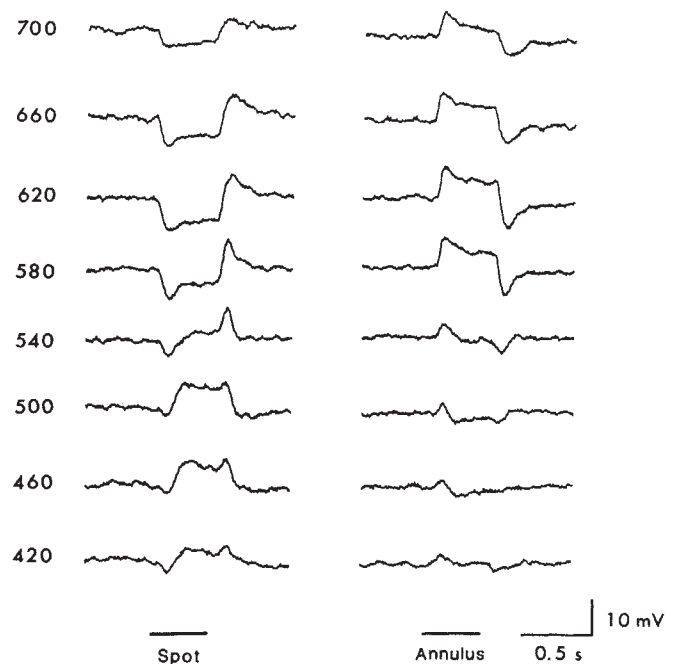


Fig. 1 Left column: receptive field centre responses of an off-centre double colour-opponent bipolar cell to wavelengths indicated (in nm) in the far left column. Stimulus diameter, $100\ \mu\text{m}$. The diameter of the RF centre, as determined from the area-amplitude curve, was $\sim 400\ \mu\text{m}$ regardless of stimulus wavelength. Light intensity was $1.6 \times 10^{12}\ \text{quanta cm}^{-2}\ \text{s}^{-1}$ at each wavelength, which corresponded to ~ 1 log unit above threshold at $620\ \text{nm}$. Right column: responses to annuli ($0.5\ \text{mm i.d.}$, $3.5\ \text{mm o.d.}$) whose intensity was $2.2 \times 10^{12}\ \text{quanta cm}^{-2}\ \text{s}^{-1}$, which was also ~ 1 log unit above threshold at $620\ \text{nm}$. Horizontal bars beneath each column of records indicate timing of stimulus; stimulation duration = $400\ \text{ms}$.

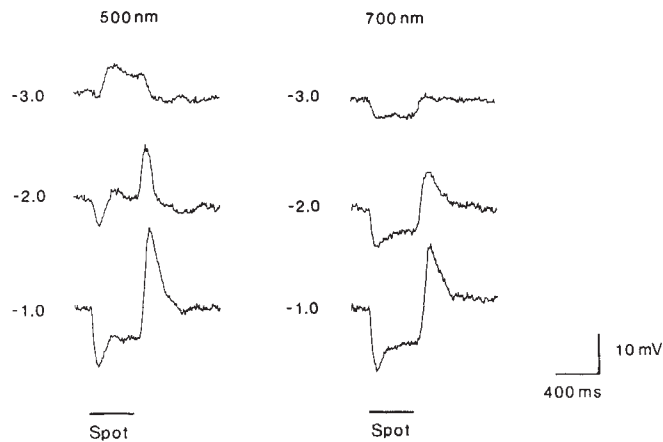


Fig. 2 Receptive field centre responses of the same cell as that shown in Fig. 1 to 500-nm spots (left column) and 700-nm spots (right column) of various intensities. Spot diameter, 100 μm . Light intensity at -3.0 log units corresponds to 9×10^{11} quanta $\text{cm}^{-2} \text{s}^{-1}$ for both wavelengths. Horizontal bars beneath each column of records indicate timing of stimulus.

light microscope examination of 25 cells which had been filled with Procion yellow¹⁰ by intracellular, iontophoretic injection. Those cells which depolarized in response to a white spot (100 μm diameter) positioned at the receptive field ('RF') centre will be referred to here as 'on-centre' cells, while those which hyperpolarized to the same stimulus are termed 'off-centre' cells. Types of cone input to both the centre- and surround-components of the RF of each bipolar cell were studied by measuring their spectral sensitivities.

Fifteen on-centre and three off-centre cells had a double colour-opponent RF. Figure 1 shows an example of an off-centre cell. The polarity of the RF centre response (left column) was hyperpolarizing to long wavelengths (580–740 nm) and depolarizing to short wavelengths (400–540 nm). Centre responses displayed transient components, hyperpolarizing at light increment and depolarizing at light decrement. At intermediate wavelengths (540–580 nm) these transients were particularly prominent. The latency of the centre hyperpolarizing (C-) component was shorter than that of the centre depolarizing (C+) component (by ~ 50 ms, at the intensity specified in Fig. 1). The C- component might also decay faster than the C+ component.

Response polarities depended not only on stimulus wavelength but also on the light intensity. Figure 2 shows responses of the same bipolar cell as in Fig. 1, to spots of long and short wavelength at various intensities. When the intensity of the 500-nm spot was increased to greater than 2 log units above threshold, the polarity of the plateau phase of the RF centre response reversed from depolarization to hyperpolarization. This may reflect a larger maximum amplitude for C- components than C+ components, or perhaps a suppression of C+ components by some unknown mechanism. In the presence of a strong 660-nm adapting spot, we were unable to elicit C+ components.

The right column of Fig. 1 shows the surround responses, that is, those elicited by annuli, at various wavelengths. At each wavelength, the polarity of the surround response was opposite to that of the centre response: depolarizing to long and hyperpolarizing to short wavelengths. The surround response had depolarizing-on and hyperpolarizing-off transients. The latency of hyperpolarizations elicited by surround stimulation (S-) was ~ 70 ms longer than that of surround depolarization (S+). In response to annuli which were more than 0.5 log units brighter than those used in the experiment of Fig. 1, the plateau phase of the surround response to any wavelength was always depolarizing. This could have resulted from an interaction between S+ and S- components, similar to that described between C- and C+ components. Centre and surround components responded

in a double opponent manner within 1.5 log units above threshold.

Figure 3 illustrates the spectral sensitivity of the RF centre of the cell described above. Above 620 nm, the spectral sensitivity agrees with those of carp red-sensitive cones (ref. 11 and A.K. and M.T., unpublished data), whereas around 500 nm, the sensitivity of this unit is nearly 0.5 log units lower. Low sensitivity in this region of the spectrum may be explained by an opponent C+ component (see below).

The spectral sensitivity of the C+ component (Fig. 3, filled circles) peaked at about 500 nm and could reflect input driven by green-sensitive cones or rods. Although the absorption spectra of rods and green-sensitive cones peak at similar wavelengths, we feel that rods are unlikely to be the origin of the C+ component for three reasons: (1) the polarity of rod signals is identical to that of red-sensitive cone signals in carp bipolar cells^{12,13}, (2) the lowest light intensity which elicited C+ and C- components in our experimental conditions differed by 0.5 log units or less, and (3) the background used in the present experiments would have reduced rod input if any was present¹⁴.

Unfortunately, the recording from the cell illustrated in Fig. 1 became unstable before the spectral sensitivity of the RF surround could be measured completely. Nonetheless, the response polarity of the plateau phase of the surround clearly indicated

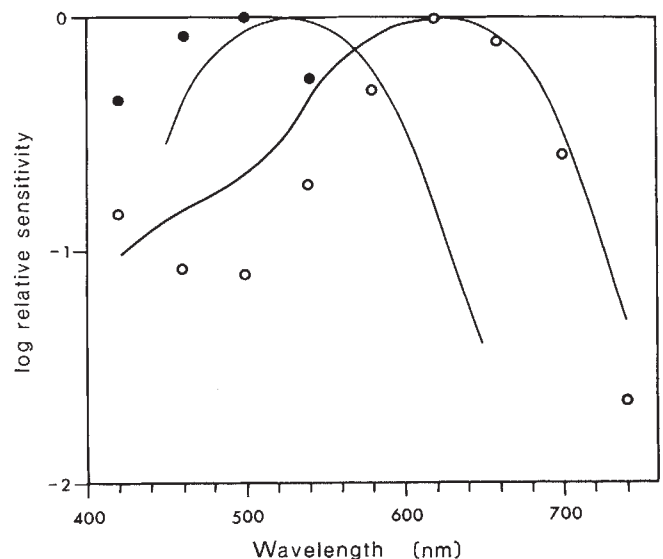


Fig. 3 Spectral sensitivity of the unit shown in Fig. 1. The sensitivity of the centre-hyperpolarizing (C-) component (open circles) was defined as the reciprocal of the light intensity necessary to elicit a 4.4-mV hyperpolarization (30% of maximum amplitude) during the plateau phase (in this measurement, at 233 ms after the stimulus onset). A similar spectral sensitivity was obtained at any time between 200 and 400 ms after the onset of the flash, but it shifted towards longer wavelengths (≤ 620 nm) when measured sooner, and towards shorter wavelengths (≥ 520 nm) if measured within 200 ms of termination of the stimulus. Short wavelengths could hyperpolarize the cell in the receptive field centre only at relatively high light intensities. The spectral sensitivity of the centre-depolarizing (C+) component (filled circles) was determined by setting the criterion at 1.6 mV depolarization from the dark membrane potential, again measuring response amplitudes at 233 ms after stimulus onset. This curve was normalized to that of the C- component by shifting it along the sensitivity axis by 0.5 log units. Continuous curve peaking at 620 nm indicates the spectral sensitivity of red-sensitive cones in the carp (A.K. and M.T., unpublished observation). The continuous curve peaking at 530 nm is the spectral absorption of green-sensitive cones¹⁵. The wavelength at which the C+ component peaks is roughly 30 nm shorter than the absorption spectral maximum of green-sensitive cones, which may be due to incomplete isolation of the C+ from the C- components. The difference between normalized spectral sensitivities of the C- component of this bipolar cell and that of red-sensitive cones can be matched with the spectral sensitivity of the C+ component.

colour-opponent input from red- and green-sensitive cones. Surround depolarization was largest at about 620 nm, whereas surround hyperpolarization was maximum at around 500 nm.

The response polarities of on-centre bipolar cells were opposite to those seen in the off-centre cell described above. The responses of these cells displayed C+ and S- components which were maximally sensitive to ~620-nm stimuli, and C- and S+ components which were maximally sensitive to intermediate wavelengths (~500 nm).

In many bipolar cells, colour opponency was less evident in the surround than in the centre. This may be partly due to the fact that the green-sensitive component in the surround saturates at lower light intensities than in the centre. As it is difficult to maintain stable penetrations of bipolar cells long enough to measure spectral sensitivities fully, the surround green-sensitive component in these cells can be easily overlooked.

The present findings indicate that at least some bipolar cells may have RF organization similar to that of double colour-opponent ganglion cells^{1-3,5}. The specific retinal circuitry which produces the response components described here are unclear. However, it seems possible that this type of bipolar cell could provide direct input to double colour-opponent retinal ganglion cells^{9,13}.

We thank Dr A. T. Ishida for helping to improve the manuscript. This work was supported in part by grants from the Ministry of Education, Science and Culture and from the Naito Science Foundation.

Received 12 May; accepted 14 July 1981.

1. Daw, N. W. *Science* **158**, 942-944 (1968).
2. Daw, N. W. *J. Physiol., Lond.* **197**, 567-592 (1968).
3. Raynaud, J.-P. *Science* **177**, 84-85 (1972).
4. Bowling, D. B. *J. Physiol., Lond.* **299**, 173-196 (1980).
5. Adams, A. J. & Afanador, A. J. *Am. J. Optom.* **48**, 889-896 (1971).
6. Michael, C. R. *Brain Behav. Evol.* **3**, 205-209 (1970); *J. Neurophysiol.* **41**, 572-588, 1233-1249, 1250-1266 (1978).
7. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **195**, 215-243 (1968).
8. Daw, N. W. *Physiol. Rev.* **53**, 571-611 (1973).
9. Famiglietti, E. V. Jr, Kaneko, A. & Tachibana, M. *Science* **198**, 1267-1269 (1977).
10. Stretton, A. O. W. & Kravitz, E. A. *Science* **162**, 132-134 (1968).
11. Tomita, T., Kaneko, A., Murakami, M. & Pautler, E. L. *Vision Res.* **7**, 519-531 (1967).
12. Kaneko, A. & Tachibana, M. *Sensory Process* **2**, 383-387 (1978).
13. Kaneko, A., Famiglietti, E. V. Jr & Tachibana, M. in *Neurobiology of Chemical Transmission* (eds Otsuka, M. & Hall, Z.) 235-251 (Wiley, New York, 1979).
14. Kaneko, A. & Yamada, M. *J. Physiol., Lond.* **227**, 261-273 (1972).
15. Hárosi, F. I. & MacNichol, E. F. Jr *J. gen. Physiol.* **63**, 279-304 (1974).

A synthetic peptide as an antagonist of substance P

Göran Engberg*, Torgny H. Svensson*†, Sune Rosell‡ & Karl Folkers§

* Department of Pharmacology, University of Göteborg, Box 33031, S-400 33 Göteborg 33, Sweden

‡ Department of Pharmacology, Karolinska Institute, S-104 01 Stockholm, Sweden

§ Institute for Biochemical Research, The University of Texas at Austin, Texas 78712, USA

The undecapeptide substance P (SP) is believed to participate in synaptic transmission in primary sensory neurones and in many other parts of the central nervous system^{1,2}, including (in the rat) the pontine structure locus coeruleus (LC). So far, studies of the role of SP in the central nervous system have been impeded by the lack of a specific antagonist to SP. Immunoneutralization of SP has been used³, and more recently synthetic analogues of SP, in which D-amino acids are substituted for L-amino acids, have been found to block peripheral effects of SP^{4,5}. We now describe experiments which show that the peptide which differs from SP in the presence of D-Pro at position 2 and D-Trp at positions 7 and 9 specifically blocks *in vivo* the SP-induced excitation of LC neurones and is thus a CNS antagonist of SP.

† To whom correspondence should be addressed.

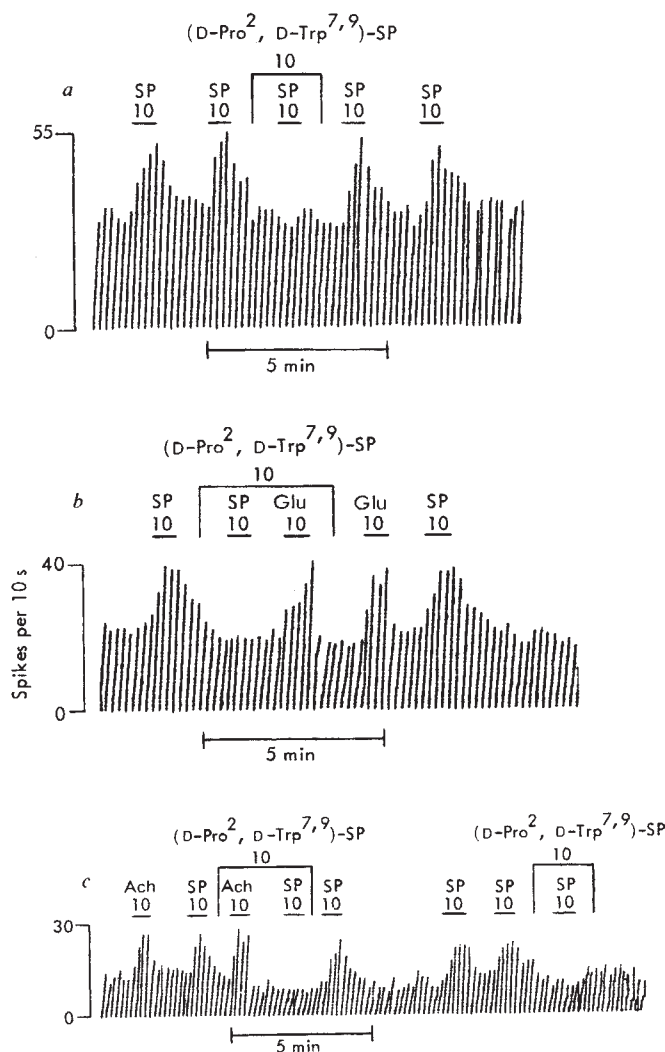


Fig. 1 The effect of microiontophoretically applied (D-Pro², D-Trp^{7,9})-SP on the excitatory responses of single LC neurones to substance P (SP), acetylcholine (ACh) and glutamate (Glu). *a*, Typical response of an LC neurone to repeated administration of SP. No tachyphylaxis is seen and the effect is blocked by simultaneous application of (D-Pro², D-Trp^{7,9})-SP. The effect of the SP antagonist is short lived as application of SP after termination of antagonist application produces a normal response. In *b* and *c* the specificity of the SP-antagonistic action of (D-Pro², D-Trp^{7,9})-SP is established. Application of the antagonist on the same neurone completely blocked the SP-induced excitation but not the typical, and probably nonspecific, excitation by glutamate (*b*). For another LC neurone (*c*), application of the antagonist again completely blocked the SP effect, without affecting the typical ACh-induced excitatory response. The records show average rate histograms, the counter being reset every 10 s. Horizontal bars indicate periods of ejection and the numbers above show ejection currents in nA.

In this study, single cell recording and microiontophoretic techniques were used to establish *in vivo* whether the peptide (D-Pro², D-Trp^{7,9})-SP would specifically block the excitation of LC neurones by SP. Male Sprague-Dawley rats (~250 g) were anaesthetized with chloral hydrate (400 mg per kg intraperitoneally) and mounted in a stereotaxic apparatus. The body temperature of the animals was kept at 36-37 °C by means of a heating pad. For single LC cell recording, a 3-mm burr hole was drilled in the skull with its centre located 1.1 mm lateral to the midline and 1.1 mm posterior to lambda. A tentative LC neurone was localized as described previously⁶⁻⁹. The final recording site was marked at the end of each experiment by iontophoretic ejection of Fast green. The rats were then perfused through the heart with 10% formaldehyde solution and

serial 50- μ m frozen sections of the brain were cut, mounted and stained with cresyl violet and counterstained with neutral red. Only units within the LC were studied.

The electrophysiological experiments were performed essentially as described previously^{10,11}. The five-barrel micropipettes used were broken to a tip diameter of 3–5 μ m and each of the barrels contained a few strands of fibre glass before pulling to facilitate filling of the tips by capillary action¹². The central barrel, filled with 2 M NaCl solution saturated with Fast green, was used to record action potentials. One of the side barrels contained 4 M NaCl solution for automatic current balancing and the others contained SP (3.0 mM), (D-Pro², D-Trp^{7,9})-SP (3.0 mM), acetylcholine chloride (ACh; 0.2 M, dissolved in 0.2 M NaCl, pH 6.2) and L-glutamic acid (20 mM, dissolved in 0.2 M NaCl, pH 8.6). Both peptides were dissolved in 20 mM sodium acetate, pH 4.5 (ref. 13).

The *in vitro* impedances, measured at 135 Hz, were typically 2.5–5 M Ω in the central barrel and 20–100 M Ω in the barrels containing drugs. The retaining currents for SP and (D-Pro², D-Trp^{7,9})-SP were 0.5 nA and 2 nA, respectively; previous experiments have shown that higher retaining currents for SP cause a considerable delay in onset of neuronal response during iontophoretic experiments¹⁴. For glutamate and ACh a retaining current of 10 nA was used. Ejections were performed frequently to avoid dead-time artefacts. The electrode potentials were passed through a high input-impedance amplifier and filters. Each spike was discriminated and fed into an integrator, which was reset every 1 or 10 s, and then displayed on an oscillographic recorder.

Most LC neurones responded with an increased firing rate to microiontophoretic application of SP at low ejection currents (Fig. 1, Table 1). The effect was rapid in onset with a delay of ~5 s. Total recovery was obtained within ~1 min. Microiontophoretic application of (D-Pro², D-Trp^{7,9})-SP in a dose that did not by itself affect the firing rate of most cells (Fig. 1, Table 1), almost completely blocked the excitatory effect of SP on LC neurones (in 22 of 23 cells). Furthermore, our data (Fig. 1) demonstrate that the SP-blocking action is short lived. The peptides did not cause any obvious alteration in spike amplitude. A few cells (6 of 20) responded to application of (D-Pro², D-Trp^{7,9})-SP by a slight decrease in firing rate. However, this did not seem to be dose-dependent—in those cases in which inhibition was obtained, the ejection current varied between 5 and 25 nA but in other cases ejection currents up to 50 nA produced no inhibition of firing. The same amounts of the peptide that blocked the excitation by SP, for the same cell did not antagonize the excitatory action of ACh or glutamate on LC neuronal firing rate (Fig. 1, Table 1). The well-known response of LC neurones to a noxious stimulus, for example, pinching of the contralateral paw, which is characterized by a burst of activity followed by a quiescent interval¹⁵, was not substantially blocked by microiontophoretic application of (D-Pro², D-Trp^{7,9})-SP to LC neurones.

These data confirm the previously observed, excitatory effect of substance P on LC neurones^{10,16}, in that, in optimal iontophoretic conditions (a low retaining current, SP dissolved in sodium acetate solution¹⁴), the effect of SP was rapid in onset—this contrasts with the previously reported delayed action of the peptide¹⁷. As reported elsewhere¹³, no tachyphylaxis to the excitatory effect of SP on LC neurones was seen. Previous data indicate that the effect of SP can probably not be ascribed to interaction with the excitatory, cholinergic input to the LC^{11,16}, or with opioid mechanisms within this nucleus^{16,17}; our results lend additional support to the former view.

The major finding of this report, that (D-Pro², D-Trp^{7,9})-SP completely blocks the effect of SP on LC neurones, even in a dose which produced no effect on its own, suggests *per se* that this compound has SP antagonistic properties. This is also supported by the fact that the same amount of (D-Pro², D-Trp^{7,9})-SP did not block the probably nonspecific, glutamate-induced excitation¹⁸ or the specific ACh-induced stimulation of LC neurones^{11,16}. This finding is in contrast to the action of the

Table 1 Incidence of responses of LC neurones to microiontophoretic application of substance P, (D-Pro², D-Trp^{7,9})-SP, glutamate and ACh

	Excitation	No change	Inhibition
Substance P (3–15)	29	1	0
(D-Pro ² , D-Trp ^{7,9})-SP (5–50)	0	14	6
Glutamate (2.5–20)	9	0	0
ACh (10–15)	7	0	0
Substance P + (D-Pro ² , D-Trp ^{7,9})-SP	1	22	0
ACh + (D-Pro ² , D-Trp ^{7,9})-SP	9	1	0
Glutamate + (D-Pro ² , D-Trp ^{7,9})-SP	9	0	0

Numbers in parentheses indicate ejection currents in nA.

presumed SP antagonist baclofen, which antagonized SP-induced excitation of LC neurones, but also consistently antagonized the excitatory effects of ACh and glutamate on these cells¹⁶. Thus, (D-Pro², D-Trp^{7,9})-SP seems to have some specificity of action.

The finding that (D-Pro², D-Trp^{7,9})-SP alone did not reduce the firing rate of most of the LC neurones tested, principally indicates that the excitatory SP input to the LC is not critical for maintenance of the spontaneous activity of the noradrenergic cells. In this regard, the SP input may be analogous to the excitatory, cholinergic input to the LC, as neither systemic nor microiontophoretic application of cholinergic antagonists reduces the spontaneous firing rate of these neurones^{12,17}. Previous studies have revealed that naloxone, when given alone, does not alter the firing rate of LC neurones¹⁶ in spite of their endogenous opioid input^{7,13,19,20}. Thus, it may be that both peptide systems are not normally active tonically, but this remains to be clarified. For example, the use of chloral hydrate anaesthesia may confound the picture, as previous experiments have revealed a reduction of SP in the brain during anaesthesia²¹. Furthermore, immunoneutralization of SP released in the substantia nigra results in inhibition of the release of ³H-dopamine (DA) from nerve terminals in the caudate nucleus³. This result indicates a tonic, excitatory SP influence on the nigro-neostriatal DA pathway. Definite conclusions regarding the physiological significance of the SP input to the LC must await further studies. However, our observations indicate that the excitation of LC neurones by noxious stimuli is probably not mediated by this SP input.

In conclusion, our data demonstrate *in vivo* that (D-Pro², D-Trp^{7,9})-SP effectively, and probably specifically, antagonizes the excitatory effect of SP on single cells in a brain nucleus, receiving a verified SP input.

We thank the Swedish MRC (grants 4747 and 4495) and Robert A. Welch Foundation for supporting this research.

Received 9 June; accepted 21 July 1981.

- Otsuka, M. & Konishi, S. *Cold Spring Harb. Symp. quant. Biol.* **40**, 135–143 (1976).
- Lembeck, F. & Zetler, G. *Int. Rev. Neurobiol.* **4**, 159–215 (1962).
- Cheramy, A., Nieoullon, A., Michelof, R. & Glowinski, J. *Neurosci. Lett.* **4**, 105–109 (1977).
- Folkers, K., Hörg, J., Rosell, S. & Björkroth, U. *Acta physiol. scand.* (in the press).
- Rosell, S. *et al. Acta physiol. scand.* (in the press).
- Korfi, J., Bunney, B. S. & Aghajanian, G. K. *Eur. J. Pharmac.* **25**, 165–169 (1974).
- Young, W. S., Bird, S. J. & Kuhar, M. J. *Brain Res.* **129**, 366–370 (1977).
- Aghajanian, G. K., Cedarbaum, J. M. & Wang, R. Y. *Brain Res.* **136**, 570–577 (1977).
- Aghajanian, G. K. *Nature* **276**, 186–199 (1978).
- Svensson, T. H. & Usdin, T. *Science* **202**, 1089–1091.
- Engberg, G. & Svensson, T. H. *J. neural Transm.* **49**, 137–150 (1980).
- Tasaki, K., Tsukuhara, U., Ito, S., Wayner, M. J. & Yu, W. Y. *Physiol. Behav.* **3**, 1009–1010 (1968).
- Guyenet, P. G. & Aghajanian, G. K. *Brain Res.* **136**, 178–184 (1977).
- Guyenet, P. G., Mroz, E. A., Aghajanian, G. K. & Leeman, S. E. *Neuropharmacology* **18**, 553–558 (1979).
- Cedarbaum, J. M. & Aghajanian, G. K. *Life Sci.* **23**, 1383–1392 (1978).
- Guyenet, P. G. & Aghajanian, G. K. *Eur. J. Pharmac.* **53**, 319–328 (1979).
- Krnjević, K. in *Substance P* (eds von Euler, U. S. & Pernow, B.) 217 (Raven, New York, 1977).
- Krnjević, K. *Physiol. Rev.* **54**, 418–540 (1974).
- Bird, S. & Kuhar, M. J. *Brain Res.* **122**, 523–533 (1977).
- Strahlendorf, H. K., Strahlendorf, J. C. & Barnes, C. D. *Brain Res.* **191**, 284–288 (1980).
- Zetler, G. & Ohnesorge, G. *Arch. exp. Path. Pharmac.* **231**, 199–210 (1957).

Cartilage proteoglycans inhibit fibronectin-mediated adhesion

A. M. Rich, E. Pearlstein, G. Weissmann
& S. T. Hoffstein*

Departments of Medicine and Pathology, New York University
School of Medicine, 550 First Avenue, New York 10016, USA

Normal tissues and organs show, on histological examination, a pattern of cellular and acellular zones that is characteristic and unique for each organ or tissue. This pattern is maintained in health but is sometimes destroyed by disease. For example, in mobile joints, the articular surfaces consist of relatively acellular hyaline cartilage, and the joint space is enclosed by a capsule of loose connective tissue with a lining of fibroblasts and macrophages. In the normal joint these cells are confined to the synovial lining and the articular surface remains acellular. In *in vitro* culture, macrophages and their precursor monocytes are very adhesive, and fibroblasts can migrate and overgrow surfaces such as collagen or plastic used for tissue culture. The fibroblasts adhere to collagen by means of fibronectin, which they synthesize and secrete¹. Because the collagen of cartilage is capable of binding serum fibronectin² and fibronectin is present in cartilage during its development³, these cells should, in theory, slowly migrate from the synovial lining to the articular surface. It is their absence from the articular cartilage in normal circumstances, and their presence in such pathological states as rheumatoid arthritis, that is striking. We therefore set out to determine whether a component of cartilage could prevent fibroblast adherence in a defined adhesion assay. As normal cartilage is composed of 50% proteoglycans and 50% collagen by dry weight⁴, we tested the possibility that the proteoglycans in cartilage inhibit fibroblast adhesion to collagen. We present here evidence that fibroblast spreading and adhesion to collagenous substrates is inhibited by cartilage proteoglycans.

Fibroblast adherence to collagen-coated dishes was measured in the presence or absence of fibronectin. As expected, only 4% of the fibroblasts adhered to collagen without fibronectin whereas 70–80% adhered after fibronectin had been bound to the collagen. When collagen was incubated with proteoglycans either before or after addition of fibronectin, fibroblast adhesion was inhibited by 60–70% (Table 1).

Clusters of chondroitin sulphate chains are released during the degradation of cartilage proteoglycans in rheumatoid arthritis and these, like the molecule from which they are derived, possess multiple anionic charges. We therefore determined whether the glycosaminoglycan, chondroitin sulphate, mimicked the inhibition of fibronectin-mediated adhesion of fibroblasts to collagen seen with intact proteoglycans. However, unlike the latter, chondroitin sulphate had no significant effect on fibronectin-mediated adhesion.

Fibronectin also mediates fibroblast spreading on dishes used for tissue culture, although adhesion to this surface is independent of fibronectin⁵. Cartilage proteoglycans also inhibited this fibroblast spreading. After 30 min, fibroblasts were well spread on plastic surfaces coated with fibronectin (92% spread; 8% round) (Fig. 1a). On wells first coated with fibronectin and then with proteoglycans, less than 1% of the cells were spread and 99% were round (Fig. 1b). In each case, 400 cells were counted.

Thus, two fibronectin-mediated processes, adhesion to collagen and spreading on plastic dishes used in tissue culture, were inhibited by the addition of cartilage proteoglycans to either substrate. These data suggested that proteoglycans bound to collagen and plastic surfaces but did not allow conclusions to be made concerning the mechanism by which proteoglycans inhibited fibronectin-mediated adhesion. To test the possibility that proteoglycans inhibited fibronectin-mediated adhesion to

collagen by preventing fibronectin binding to surfaces or by displacing previously bound fibronectin, we examined the binding of iodinated proteoglycans and fibronectin to collagen-coated surfaces. The interactions of proteoglycans and collagen were studied by incubating ¹²⁵I-labelled proteoglycans with collagen-coated wells. Because preliminary studies indicated that collagen bound large quantities of proteoglycans, the amount of collagen per well was reduced to 10 µg. Otherwise, the conditions duplicated those in the cell adherence studies. As increasing concentrations of proteoglycans were added to collagen-coated wells, the amount bound increased linearly up to a concentration of 20 µg per well (Fig. 2). Thus, the binding of the proteoglycans to collagen was saturable and saturation was reached at 20 µg proteoglycans per 10 µg collagen per well. In contrast, the nonspecific binding of proteoglycans to plastic was linear and not saturable. Extrapolating from this data, saturation of 500 µg of collagen (the concentration used in the cell adherence studies) would occur at a concentration of ~1 mg proteoglycans per well or 3.3 mg proteoglycans ml⁻¹. The fibroblast adhesion studies were done at a proteoglycan concentration of 5 mg ml⁻¹, a concentration above saturation.

Binding was specific as well as saturable as it could be blocked by a 400-fold excess of unlabelled proteoglycans or chondroitin sulphate, but not by bovine serum albumin (BSA) (Table 2). Chondroitin sulphate was not as effective at blocking labelled proteoglycan binding as was intact cartilage proteoglycan. The fact that chondroitin sulphate blocked the binding of proteoglycan to collagen suggests that the proteoglycan-collagen interactions were due, at least in part, to electrostatic interactions.

Having shown that cartilage proteoglycans bind to collagen, we next studied whether proteoglycans inhibited binding of

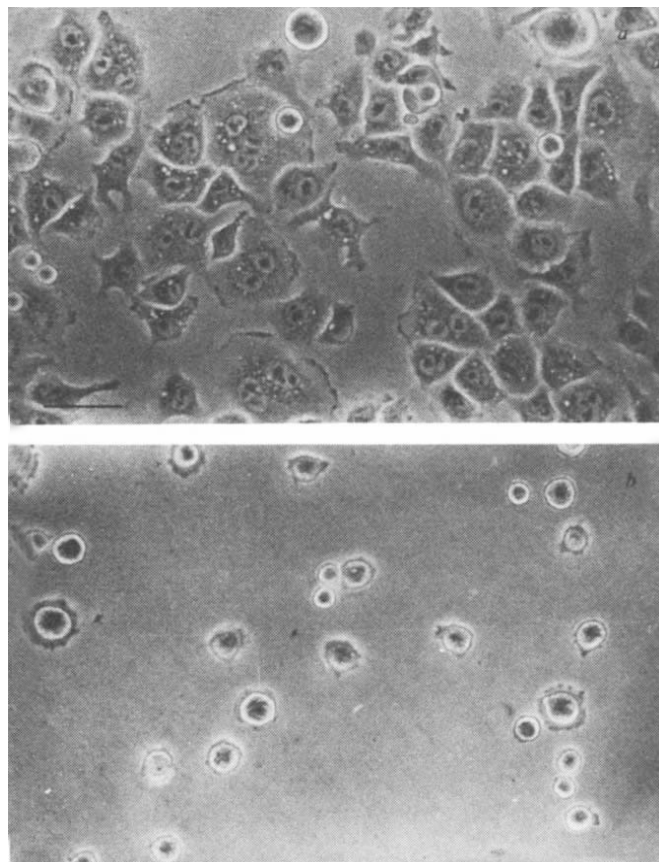


Fig. 1 Spreading of fibroblasts on plastic coated with fibronectin or with fibronectin and proteoglycans. Tissue culture plastic wells (Falcon) were incubated with fibronectin (30 µg ml⁻¹) for 30 min at 37 °C and then with either PBS (a) or proteoglycans (5 mg ml⁻¹) (b), for 30 min at 37 °C in a CO₂ atmosphere. Cells were photographed using an Olympus IMT inverted microscope and an Olympus PM-10AD 35 mm camera system. Scale bar, 50 µm.

* To whom reprint requests should be addressed.

Table 1 Proteoglycan monomers inhibit fibronectin-mediated adhesion of PyBHK fibroblasts to collagen

Plate incubation		% Control adhesion ± s.e.m. (n)
First	Second	
Fibronectin	PBS	100
PBS	Fibronectin	100
Fibronectin	Proteoglycans	34.3 ± 6.8 (3)
Proteoglycans	Fibronectin	36.3 ± 3.8 (3)
PBS	Fibronectin and proteoglycans	28.3 ± 2.9 (3)
PBS	Fibronectin and chondroitin sulphate	86.0 ± 3.5 (3)
PBS	PBS	3.3 ± 0.3 (3)
PBS	Proteoglycans	9.0 ± 1.9 (4)
PBS	Chondroitin sulphate	13.3 ± 4.0 (3)

Proteoglycan monomers inhibit fibronectin-mediated adhesion of PyBHK fibroblasts to collagen. Collagen (Sigma type III; actually type I collagen) was prepared as described elsewhere¹⁰ and 300 µl were plated in to wells of flexible, clear vinyl, flat-bottomed 96-well tissue culture plates (growth area of well = 1.67 cm²; Linbro). Fibronectin was isolated from human serum as previously described¹¹ and diluted to 30 µg ml⁻¹ in phosphate-buffered saline (PBS) containing Mg²⁺ and Ca²⁺. Cartilage proteoglycan monomers were prepared from bovine nasal septal cartilage (supplied by Dr L. Rosenberg)¹². Proteoglycans were dissolved in PBS at 5 mg ml⁻¹. Each incubation was for 30 min at 37 °C and wells were washed once with PBS after each incubation. So that added fibronectin was the only source of fibronectin in the system, we used freshly trypsinized PyBHK fibroblasts (ICRF), a transformed cell line that synthesizes fibronectin at a relatively low level¹³. These cells, in Dulbecco's modified Eagle's medium (DMEM; Gibco), were incubated in the wells (20,000 per well) for 90 min at 37 °C in a CO₂ atmosphere. The wells were washed three times with PBS and the adherent cells removed with trypsin-EDTA and counted in a haemocytometer. Data are expressed as % cells that adhered ± s.e.m. (n = number of experiments) compared with the number of cells that adhered when only fibronectin was present (15,120 ± 1,230; n = 5).

fibronectin to collagen or displaced fibronectin that was already bound. Collagen-coated wells were incubated with proteoglycans before and after incubation with ¹²⁵I-fibronectin (Table 3). Previous binding of proteoglycans to collagen neither inhibited nor enhanced the binding of fibronectin to collagen and the

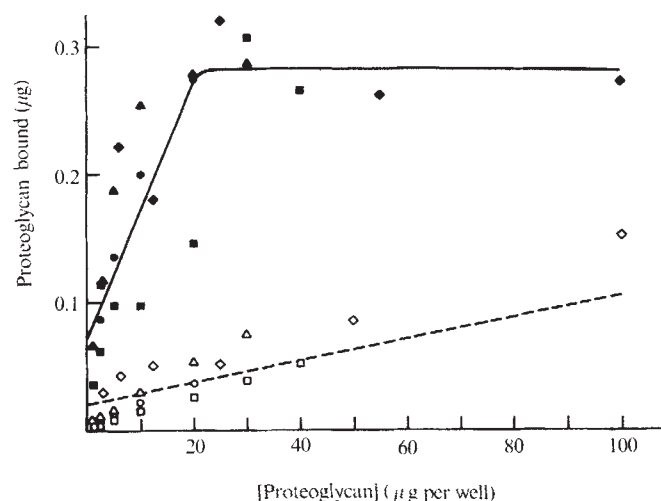


Fig. 2 The binding of increasing concentration of proteoglycans to collagen and plastic. Collagen was prepared as described in Table 1 legend and diluted 50-fold before being placed in wells. The wells were rinsed once before use. Proteoglycans were iodinated using the chloramine T method¹⁶. A mixture of one part ¹²⁵I-proteoglycan to 20 parts nonradioactive proteoglycan in 1% BSA (Sigma) in PBS was incubated in both plastic (open symbols) and collagen-coated wells (closed symbols) for 30 min at 37 °C. Wells were washed three times with PBS and the flat portion of the wells removed with a razor blade and counted. Four separate experiments (Δ, □, ◇, ○) were done and each point is the average of three wells.

incubation of collagen-coated wells with proteoglycans after binding of fibronectin to the collagen did not displace the fibronectin that was bound. These two lines of evidence suggested that proteoglycans and fibronectin were bound to different sites on collagen. To determine whether cartilage proteoglycans bound to fibronectin, we incubated plastic wells with either 1% BSA or fibronectin (30 µg ml⁻¹) followed by 1% BSA, and then with ¹²⁵I-proteoglycan (70 µg ml⁻¹). Plastic wells coated with fibronectin bound 2.84 ng ± 0.29 proteoglycan (mean ± s.e.m. of three experiments). This is more than twice the amount that bound to BSA-coated plastic wells (1.24 ng ± 0.25). These experiments indicate that proteoglycans do bind to fibronectin, but not to surfaces coated with other proteins (for example, BSA).

Our data show that cartilage proteoglycans bind to type I collagen in a specific and saturable way. Proteoglycans do not prevent fibronectin binding nor do they displace bound fibronectin. However, collagen to which both proteoglycans and fibronectin have been bound cannot support fibroblast adherence. As chondroitin sulphate did not inhibit fibroblast adhesion, inhibition of fibroblast adhesion by proteoglycans may require their binding to collagen⁶ or to fibronectin⁷ by the protein core. Alternatively, the large size⁴ of proteoglycan monomers (average molecular weight = 2.5 × 10⁶) may, by steric hindrance, mask the cell binding site of fibronectin. This hypothesis is supported by the results of Weiss and Reddi³, who found that hyaluronidase treatment of cartilage rendered endogenous fibronectin identifiable by immunofluorescence.

Knox and Wells⁸ have also found that exogenous cartilage proteoglycan monomers, but not chondroitin sulphate, inhibited chick embryo fibroblast attachment to collagen and plastic. The exact role of fibronectin is unclear from their system because chick embryo fibroblasts produce a high level of fibronectin⁹ and

Table 2 Specificity of cartilage proteoglycan binding to collagen-coated wells

	% Proteoglycan bound ± s.e.m. (n)
¹²⁵ I-proteoglycan	100
¹²⁵ I-proteoglycan in 1% BSA	86 ± 16 (2)
¹²⁵ I-proteoglycan in 0.1% chondroitin sulphate	31 ± 5 (4)
¹²⁵ I-proteoglycan in 0.1% proteoglycan	12 ± 2.2 (5)

The binding of proteoglycans to collagen is inhibited by excess chondroitin sulphate or proteoglycan. Collagen was prepared as described in Fig. 2 legend and incubated with 0.25 µg ¹²⁵I-proteoglycan per well in PBS with or without 1% BSA, 0.1% chondroitin sulphate or 0.1% proteoglycan. Data are expressed as % total proteoglycans that bound in PBS alone (0.070 µg per well ± 0.008, n = 5) ± s.e.m. (n = number of experiments).

Table 3 Effect of proteoglycans on binding of fibronectin to collagen-coated wells

Plate incubation		µg Fibronectin bound ± s.e.m. (n)
First	Second	
BSA	¹²⁵ I-fibronectin	1.44 ± 0.095 (4)
Proteoglycans	¹²⁵ I-fibronectin	1.25 ± 0.100 (4)
¹²⁵ I-fibronectin	BSA	0.69 ± 0.110 (3)
¹²⁵ I-fibronectin	Proteoglycans	0.66 ± 0.092 (3)

Proteoglycans do not inhibit the binding of fibronectin to collagen. Collagen was prepared and incubations carried out as described in Table 1 legend. The concentration of proteoglycan was 5 mg ml⁻¹; that of fibronectin, iodinated using lactoperoxidase¹⁴, was 30 µg ml⁻¹. Data are expressed as µg fibronectin bound per well ± s.e.m. (n = number of experiments). The decrease in amount of fibronectin bound when fibronectin was the first incubation is presumably due to loss of collagen from the surface that occurred during incubation at 37 °C (ref. 15).

many of their experiments were done in medium containing serum.

Our data support the hypothesis that it is the proteoglycan component of normal cartilage which prevents fibroblast adherence to the articular surface. The loss of proteoglycans from cartilage or other tissues during inflammation could create a suitable substrate for the migration of adjacent fibroblasts. Thus in rheumatoid arthritis, proteoglycan loss from cartilage may be a prerequisite for the invasion of the articular surface by synovial lining cells.

This work was supported by NIH grants AM 25776, AM 01431 (USPHS), HL 19721 and AM 11949. A.M.R. was supported by grant 5-T32CA-09161 from the NCI DHEW and NIH grant AI 17365. S.T.H. is the recipient of a Research Career Development Award (HL-00356).

Received 9 June; accepted 21 July 1981.

1. Pearlstein, E., Gold, L. I. & Garcia-Pardo, A. *Molec. cell. Biochem.* **29**, 103-128 (1980).
2. Gold, L. I. & Pearlstein, E. *Biochem. J.* **186**, 551-554 (1980).
3. Weiss, R. E. & Reddi, A. H. *J. Cell Biol.* **88**, 630-636 (1981).
4. Keiser, H. D. in *The Cell Biology of Inflammation* (ed. Weissmann, G.) (Elsevier/North-Holland, Amsterdam, 1980).
5. Grinnell, F. *Int. Rev. Cytol.* **53**, 65-144 (1978).
6. Toole, B. P. *J. Biol. Chem.* **251**, 896-897 (1976).
7. Culp, L. A., Murray, B. A. & Rollins, B. J. *J. supramolec. Struct.* **11**, 401-427 (1979).
8. Knox, P. & Wells, P. *J. Cell Sci.* **40**, 77-99 (1979).
9. Yamada, K. M. *J. Cell Biol.* **78**, 520-541 (1978).
10. Pearlstein, E. *Nature* **262**, 497-500 (1976).
11. Engvall, E. & Ruoslahti, E. *Int. J. Cancer* **20**, 1-5 (1977).
12. Rosenberg, L., Hellmann, W. & Kleinschmidt, A. K. *J. Biol. Chem.* **250**, 1877-1883 (1975).
13. Pearlstein, E. & Waterfield, M. D. *Biochim. biophys. Acta* **362**, 1-12 (1974).
14. Pearlstein, E. & Baez, L. *Analgt. Biochem.* (in the press).
15. Kleinman, H. K., McGoodwin, E. G., Rennard, S. I. & Martin, G. R. *Analyt. Biochem.* **94**, 308-312 (1979).
16. McConahey, P. J. & Dixon, F. J. *Int. Archs Allergy appl. Immun.* **29**, 185-189 (1966).

Cell-specific drug transfer from liposomes bearing monoclonal antibodies

Lee D. Leserman, Patrick Machy & Jacques Barbet

Centre d'Immunologie INSERM-CNRS de Marseille-Luminy,
Case 906, 13288 Marseille Cédex 9, France

Liposome-encapsulated drugs and fluorescent markers have been proposed as useful tools in studies of cell biology and in chemotherapy¹⁻³. Until recently, this approach has been limited by the lack of means to fix protein ligands, such as antibodies, to liposome surfaces where they can mediate specific interactions with cells. Our newly developed technique⁴ for the covalent coupling of monoclonal antibodies to small, sonicated liposomes was used to study the specific delivery of methotrexate (MTX) to cellular targets *in vitro*. Liposomes containing MTX and carboxyfluorescein (CF)⁵⁻⁷ were coupled to monoclonal antibodies against determinants of molecules encoded by the major histocompatibility complex of the murine H-2^k haplotype. These were incubated with lipopolysaccharide (LPS)-induced blast cells from CBA (H-2^k), or C57BL/6 (B6) (H-2^b) spleens. Only spleen cells from the CBA strain showed evidence of drug effect, as indicated by inhibition of radiolabelled deoxyuridine incorporation, demonstrating that the interaction of the liposomes with the cells was specific. With regard to the different determinants, the effect of the liposome-encapsulated MTX was not correlated with the number of liposomes bound to cells. More liposomes became bound to the H-2K^k molecule than to the H-2 I-E^k molecule, but liposomes bound to the latter were more effective for drug delivery. An endocytic pathway for the liposome-encapsulated drug is suggested by the fact that the effect of the drug in liposomes, but not in free form, was inhibited by the lysosomotropic amine, NH₄Cl. The techniques described here provide a simple quantitative assay for the 'endocytic potential' of any cell-surface determinant for which a ligand, such as an antibody or hormone, is available.

Murine IgG2a, κ monoclonal antibodies 11-4-1 (ref. 8) (specificity for H-2K^k) and H40.242.3, which recognizes H-2 I-E^k, were coupled to liposomes containing MTX and CF using the heterobifunctional cross-linking reagent *N*-hydroxy-succinimidyl 3-(2-pyridylthio) propionate (SPDP), as described elsewhere⁴. Aliquots of the liposome preparations were added to duplicate cultures of LPS-induced blasts from CBA or B6 mice. After incubation, the amount of cell-associated CF was determined by fluorometry after detergent lysis of cells and liposomes^{5,9}. Results are given in Table 1. Liposome binding was specific; CBA cells bound many more liposomes than did B6 cells, and the binding was inhibited by preincubation of the cells with the uncoupled homologous antibody, but not with an irrelevant antibody of the same class. The determinant recognized by the liposome-bound anti-H-2K^k antibody was richly represented on the CBA cells, the determinant recognized by the liposome-bound anti-H-2 I-E^k antibody being present in smaller amounts.

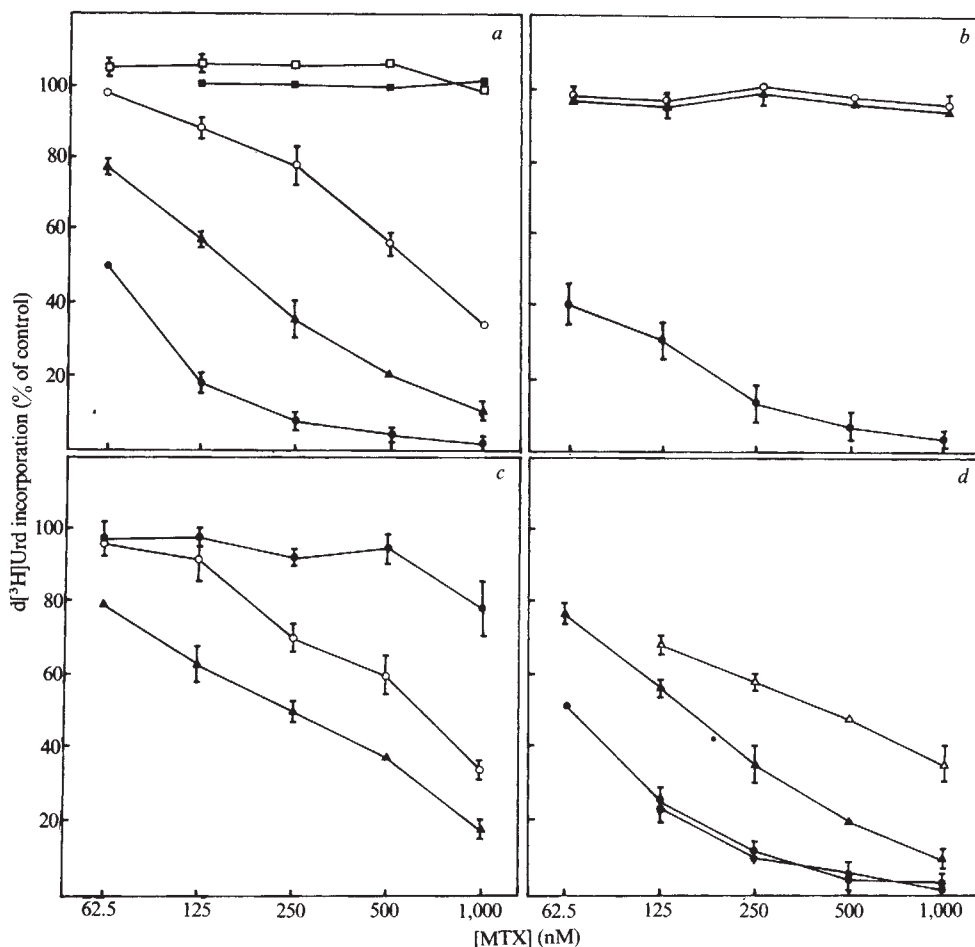
We next demonstrated that bound liposomes were capable of transferring their contents into the target cells. ³H-deoxyuridine (d[³H]Urd) uptake has been used as an index of MTX effect on tumour cells^{6,7}. Using the same technique, LPS blasts were exposed to various concentrations of MTX in free solution or in antibody-bearing liposomes. As seen in Fig. 1a, b, cells from both mouse strains were sensitive to free MTX, but only CBA cells were sensitive to MTX in liposomes bearing anti-H-2K^k or anti-H-2 I-E^k antibodies. Furthermore, the magnitude of the drug effect was not directly related to the density of the target antigen: two times more anti-H-2K^k than anti-H-2 I-E^k antibody-bearing liposomes became associated with cells (Table 1), but the number of anti-H-2K^k antibody-bearing liposomes required for 50% inhibition of d[³H]Urd incorporation was four times greater. Thus, liposomes bound to the H-2 I-E^k determinant were eight times more effective in drug delivery than

Table 1 Fluorescence of LPS blasts incubated with antibody-bearing liposomes containing carboxyfluorescein

Liposomes bearing	Free antibody in solution	Cell-associated CF (pmol) Strain	
		CBA	B6
11-4-1	None	29	0.4
11-4-1	11-4-1	0.5	ND
11-4-1	B1.1G6	25	ND
H40.242.3	None	13.7	1.2
H40.242.3	H40.242.3	0.5	ND
H40.242.3	B1.1G6	14.5	ND

Antibody B1.1G6 is IgG2a, κ and is directed against human β_2 -microglobulin. It has no detectable binding to mouse spleen cells⁴. H40.242.3 (the gift of M. Pierres) is the product of a somatic cell hybrid between A.TH spleen cells hyperimmune to A.TL and the X63-Ag8.653 myeloma. It recognizes, on H-2 I-E^k-bearing cells, a determinant analogous to the Ea chain Ia^k specificity. A detailed description of its binding properties will be reported elsewhere. All antibodies were purified on protein A-Sepharose columns¹⁴. Small, sonicated liposomes composed of 65% (molar) dipalmitoyl-phosphatidylcholine, 34% cholesterol and 1% SPDP-modified phosphatidylethanolamine⁴, containing 40 mM CF, were incubated with antibodies modified with 5 mol SPDP per mol antibody, at a final antibody concentration of 175 μ g ml⁻¹. In these conditions, 35-55% of the antibodies became liposome bound, which corresponds to averages of five to nine antibody molecules bound per liposome¹⁵. All coupling reactions resulted in at least 75% antibody-bearing liposomes, as revealed by liposome precipitation with *Staphylococcus aureus*, strain Cowan I (ref. 4). Spleen cells (30 $\times$ 10⁶) from CBA or B6 mice were incubated in tissue culture flasks with 20 ml of RPMI 1640 medium supplemented with 5% fetal calf serum (FCS), and 20 μ g ml⁻¹ of LPS from *Escherichia coli* (Difco). After 40 h incubation, cells were resuspended in the same medium without LPS at a concentration of 3 $\times$ 10⁶ cells ml⁻¹. 100- μ l aliquots were placed in flat-bottom wells of 96-well microtitre plates (Linbro). 10 μ l of medium or antibody at a final concentration of 20 μ g ml⁻¹ were added to duplicate wells. After 15 min of preincubation, 10 μ l of antibody-bearing liposomes containing a total of 220 pmol CF (lipid concentration 0.25 mM, expressed as total lipid) were added. Cells were incubated for 3 h at 37°C, after which binding had reached plateau levels. Cells were then washed three times by centrifugation. The pellets were lysed using 0.5% Triton X-100, and fluorescence measured on an Aminco SPF 500 spectrofluorometer (excitation 488 nm, emission 520 nm). Fluorescence calculations were based on a 20 nM CF standard. Quantitatively similar results were obtained in six similar experiments using five different liposome preparations. ND, not determined.

Fig. 1 To wells of microtitre plates containing 2×10^5 40-h LPS blasts in RPMI 1640 plus 5% FCS were added either free MTX or antibody-bearing liposomes containing 20 mM MTX and 40 mM CF, or 40 mM CF alone. The amount of liposome-encapsulated MTX actually used in culture was determined by measurement of the co-encapsulated CF by fluorescence after Triton lysis of liposomes. Liposome encapsulation of CF and MTX, and their leakage from liposomes, have been shown to be commensurate⁷, and we have confirmed this for antibody-bearing liposomes using radiolabelled MTX (data not shown). *a, b*, After 3 h incubation at 37 °C, 1 μ Ci of [3 H]Urd was added to cultures, and after 16 additional hours of culture, cells were collected and the concentration of incorporated [3 H]Urd measured as described elsewhere^{6,7}. *a*, CBA blasts; *b*, B6 blasts. ●, Free MTX; ▲, MTX in anti-H-2 I-E^k antibody-bearing liposomes; ○, MTX in anti-H-2K^k antibody-bearing liposomes; □, MTX in anti-human β_2 -microglobulin antibody-bearing liposomes; ■, anti-H-2 I-E^k antibody-bearing liposomes without MTX. For the representative experiment shown here, control CBA cells incorporated $102,000 \pm 2,000$ c.p.m. and B6 cells incorporated $74,000 \pm 1,400$ c.p.m. Results are the mean $\pm$ s.d. unless smaller than the points as plotted. *c*, After 3 h incubation of CBA LPS blasts with free MTX or MTX in antibody-bearing liposomes, plates were centrifuged at 1,600 r.p.m. and washed three times. They were resuspended in fresh medium and pulsed with [3 H]Urd, and then collected as above. Control cells incorporated $81,000 \pm 2,000$ c.p.m. ●, Free MTX; ▲, MTX in anti-H-2 I-E^k antibody-bearing liposomes; ○, MTX in anti-H-2K^k antibody-bearing liposomes. *d*, NH_4Cl (10 mM final concentration) was added to wells containing CBA LPS blasts 15 min before addition of free MTX or MTX in anti-H-2 I-E^k antibody-bearing liposomes. Control cells incubated with 10 mM NH_4Cl alone incorporated 82,000 c.p.m. ●, Free MTX; ★, free MTX plus NH_4Cl ; ▲, MTX in anti-H-2 I-E^k antibody-bearing liposomes; △, MTX in anti-H-2 I-E^k antibody-bearing liposomes plus NH_4Cl .



those bound to the H-2K^k determinant. Neither anti-H-2K^k nor anti-H-2 I-E^k antibody-bearing liposomes lacking MTX had any effect on d[3 H]Urd incorporation. As expected from the effect on binding shown in Table 1, anti-H-2K^k or anti-H-2 I-E^k antibodies in free solution blocked the drug effects of liposomes bearing the homologous antibody. The anti-human β_2 -microglobulin antibody in free solution had no effect on MTX delivery (data not shown).

The greater efficiency in drug delivery of anti-H-2 I-E^k antibody-bearing, as opposed to anti-H-2K^k antibody-bearing, liposomes seems to be a property of the target determinant and is not related to some peculiarity of these antibodies, as verified by the use of panel of monoclonal antibodies directed at the same molecule (data not shown).

The mechanism of entry into cells of the liposome-encapsulated drug was investigated. Cells were incubated with free MTX or with MTX in antibody-bearing liposomes for 3–4 h, then non-bound drug was washed away. As seen in Fig. 1c, the effect of the free MTX was largely eliminated, but the liposome-bound drug was as effective as in the experiment shown in Fig. 1a, in which non-bound liposomes were not washed away. As we know that the number of liposomes bound is only a fraction of those present in culture (Table 1), the drug effect cannot be due to simple leakage of MTX from cell-bound liposomes. This is further suggested by the fact that the quantitatively more important determinant is a poorer target for liposome-bound drug, and by our previous demonstration that another cell-surface determinant (surface immunoglobulin of myeloma tumour cells) did not mediate drug transfer⁹.

We⁶ and others¹⁰ have observed the apparent delivery of antibody-opsonized liposomes to lysosomes of phagocytic

tumour cells. That liposomes bound to LPS blasts, which are of B-cell origin, might also be delivered to lysosomes was suggested by an experiment using the lysosomotropic amine, NH_4Cl . When blast cells incubated with 10 mM NH_4Cl were exposed to liposomes bearing the anti-H-2 I-E^k antibody, the drug effect was markedly reduced (65–70% over a series of experiments). There was no effect of NH_4Cl on the number of cell-bound liposomes, as shown by experiments of the type described in Table 1 legend (data not shown).

A possible explanation for the selective effect of NH_4Cl on liposome-encapsulated drug is provided by the ability of NH_4Cl to neutralize lysosomal pH¹¹. When delivered into lysosomes at their normal, acidic pH, the anionic MTX would be protonated, and could thus diffuse out of lysosomes to its cytoplasmic site of action. If the lysosomal pH were neutral, this mode of entry would largely be eliminated. We do not exclude other mechanisms of action for NH_4Cl , such as inhibition of lysosomal enzymes or a direct effect on phagocytosis. It is unknown whether the apparent internalization of the H-2 I-E^k molecule is related to the physiological role of H-2 I-E determinants (or of Ia molecules in general) in the regulation of cell interactions and of antigen presentation in the immune response^{12,13}.

MTX was provided by the Drug Synthesis and Chemistry Branch, Division of Cancer Treatment, NCI. This research was supported by INSERM and CNRS institutional grants, by a contract under the programme INSERM-NCI Recherches Cliniques en Cancérologie and by a grant from Association pour le Développement de la Recherche sur le Cancer. We thank Michel Pierres and Bernard Malissen for monoclonal antibodies and our colleagues at the Centre d'Immunologie for helpful discussions.

Received 11 May; accepted 30 July 1981.

- Gregoriadis, G. (ed.) *Drug Carriers in Biology and Medicine* (Academic, London, 1979).
- Gregoriadis, G. & Allison, A. C. (eds) *Liposomes in Biological Systems* (Wiley, New York, 1980).
- Tom, B. H. & Six, H. R. (eds) *Liposomes and Immunobiology* (Elsevier, Amsterdam, 1980).
- Leserman, L. D., Barbet, J., Kourilsky, F. M. & Weinstein, J. N. *Nature* **288**, 602–604 (1980).
- Weinstein, J. N., Blumenthal, R., Sharrow, S. O. & Henkart, P. *Biochim. biophys. Acta* **509**, 272–288 (1978).
- Leserman, L. D., Weinstein, J. N., Blumenthal, R. & Terry, W. D. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4089–4093 (1980).
- Leserman, L. D., Weinstein, J. N., Moore, J. J. & Terry, W. D. *Cancer Res.* **40**, 4768–4774 (1980).
- Oi, V. T., Jones, P. P., Goding, J. W., Herzenberg, L. A. & Herzenberg, L. A. *Curr. Topics Microbiol. Immun.* **81**, 115–129 (1978).
- Leserman, L. D., Weinstein, J. N., Blumenthal, R., Sharrow, S. O. & Terry, W. D. *J. Immun.* **122**, 585–591 (1979).
- Lewis, J. T., Hafemann, D. G. & McConnell, H. M. *Biochemistry* **19**, 5376–5386 (1980).
- Ohkuma, S. & Poole, B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3327–3331 (1978).
- Martinez-A., C., Coutinho, A. & Bernabe, R. R. *Immunogenetics* **10**, 299–303 (1980).
- Sprent, J. & Alpert, B. *J. exp. Med.* **153**, 823–830 (1981).
- Goding, J. W. *J. immun. Meth.* **20**, 241–253 (1978).
- Barbet, J., Machy, P. & Leserman, L. D. *J. supramolec. Struct. cell. Biochem.* (in the press).

Genetic modification of potassium channels in *Drosophila Shaker* mutants

Lawrence Salkoff & Robert Wyman

Department of Biology, Yale University, Box 6666, New Haven, Connecticut 06511, USA

Previous research in *Drosophila* has implied that a variety of mutations affect ion channels^{1–4}. In one of these studies it was suggested that mutations mapping at the X-linked *Shaker* locus of *D. melanogaster* might cause defects in some type of potassium channel in larval motoneurons². We show here, by voltage-clamp analysis, that mutations at the *Shaker* locus affect one of the two types of voltage-sensitive potassium channels present in the *Drosophila* dorsal longitudinal flight muscle (DLM). Initial voltage-clamp studies^{5,6} revealed that one set of channels carries a fast, transient outward current, similar to the molluscan A current^{7–9}, that is altered by mutations at the *Shaker* locus. The other set of channels carries a slower activating current, similar to delayed rectification^{10,11}, that is not affected by the *Shaker* mutation.

The mutant strains used were *Sh*^{KO120} and *Sh*^{KS133} (ref. 2; from Dr Seymour Benzer), and *Sh*⁵ (ref. 12; from Dr David Suzuki). Wild-type stocks were used for comparison purposes (from Dr Donald Poulson) and included Canton-S, Oregon-R, Noumea and Canberra. The effects of the *Shaker* mutations on the potassium currents in the DLM were analysed by voltage-clamp techniques as described elsewhere^{5,6}. All experiments were performed at 4 °C. To study the A current in isolation from other currents, we selected animals at ~70 to 80 h of pupal development, a stage at which the A current is prominent but delayed rectification is small or absent⁶.

Voltage-clamp experiments showed that the *Sh*⁵ mutation greatly accelerates the rate at which the A-current channels inactivate and the rate at which they recover from inactivation. Other properties of the channels are not apparently affected. A comparison of A current in wild-type and *Sh*⁵ DLM is shown in Fig. 1; the current responses were evoked by voltage-clamp step pulses to different voltages from a holding potential of –80 mV. The A-current channels open (activate) rapidly, but close (inactivate) over a slower time course. The rapid activation of the A current can be seen in the upper traces of Fig. 1a,b as the very rapid rise (upwards) of the current to its peak. Inactivation is seen as the exponential decay (downwards) of the current from its peak. At each voltage level the inactivation of the channels occurs about twice as rapidly in *Sh*⁵ than in wild-type muscle

fibres. The current trace that falls below the baseline is the current response evoked by a 100-mV hyperpolarizing voltage-clamp step pulse; it is the same in a and b of Fig. 1, showing that the input resistance is almost identical in *Sh*⁵ and wild-type cells.

Figure 2 compares the inactivation time constants measured at different voltages for the different genotypes. The time constants for *Sh*⁵/*Sh*⁵ muscle cells (solid symbols) were about half those of wild-type muscle cells (open symbols). Four different wild-type strains were used as a control to detect any scatter due to natural variation. In no case did a data point obtained for a wild-type muscle cell overlap with one obtained for a *Sh*⁵ muscle cell.

In addition to its effect on reducing channel inactivation time, the *Sh*⁵ mutation also shortens the time required to recover from inactivation. Recovery of inactivation was investigated in a two-pulse experiment. The ratio of the peak A current during the second pulse to that during the first was plotted against the inter-pulse interval. The longer the interval, the larger the ratio; recovery followed an approximately exponential time course with time constants ($\pm$ s.d.) of 633 ± 21 ms ($N = 3$) for wild-type, 485 ± 59 ms ($N = 4$) for heterozygotes, and 351 ± 30 ms ($N = 4$) for homozygous mutants.

Although the properties of inactivation and recovery are affected by the *Sh*⁵ mutation, other properties of the channels are not obviously affected. Both wild-type and mutant cells showed a similar voltage-dependence of steady-state inactivation, with 50% inactivation at ~–42 mV. A comparison of the current-voltage relationships for the peak of the A current

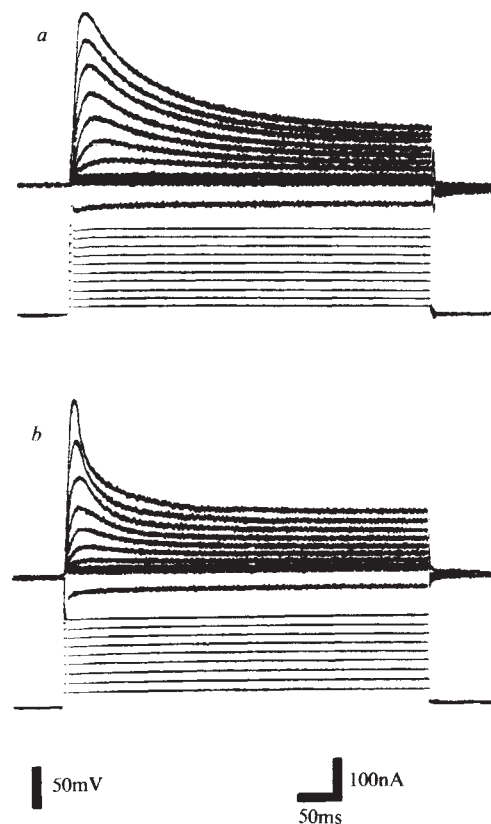


Fig. 1 Comparison of A-current responses to voltage-clamp step pulses in wild-type (a) and *Sh*⁵ (b) animals. The holding potential is –80 mV. Ten traces are shown in which single voltage-step pulses were applied to the membrane (lower traces); the membrane current response is shown in the upper traces. Outward currents are in the upward direction. Each of the current traces shows the membrane response at a different voltage level. (The voltage records for hyperpolarizing pulses to –180 mV are not shown.)

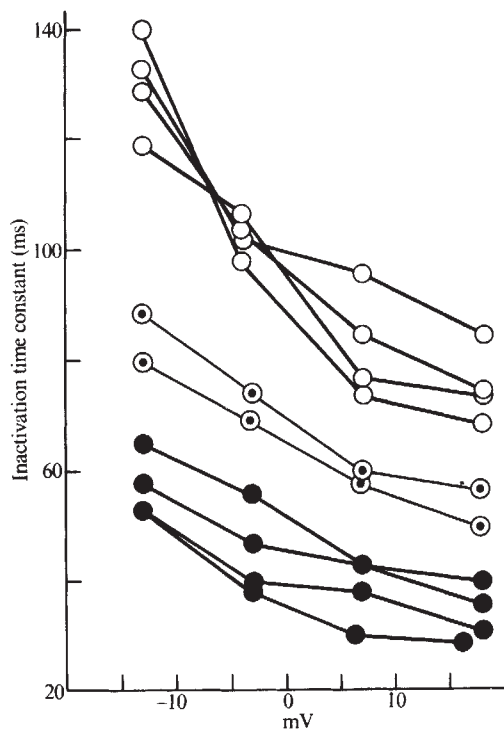


Fig. 2 Comparison of inactivation time constants at different voltages for wild-type, Sh^5 , and heterozygous (wild-type/ Sh^5) animals. The data are from experiments similar to those described in Fig. 1. The membrane voltage is shifted from a holding potential of -80 mV to the voltage indicated on the abscissa. The time constant of the exponential decay of the A current is plotted on the ordinate. $\circ$, wild-type; $\bullet$, mutant; $\odot$, heterozygote. Each wild-type curve is from a different wild-type strain and each curve is for a muscle fibre from a different animal.

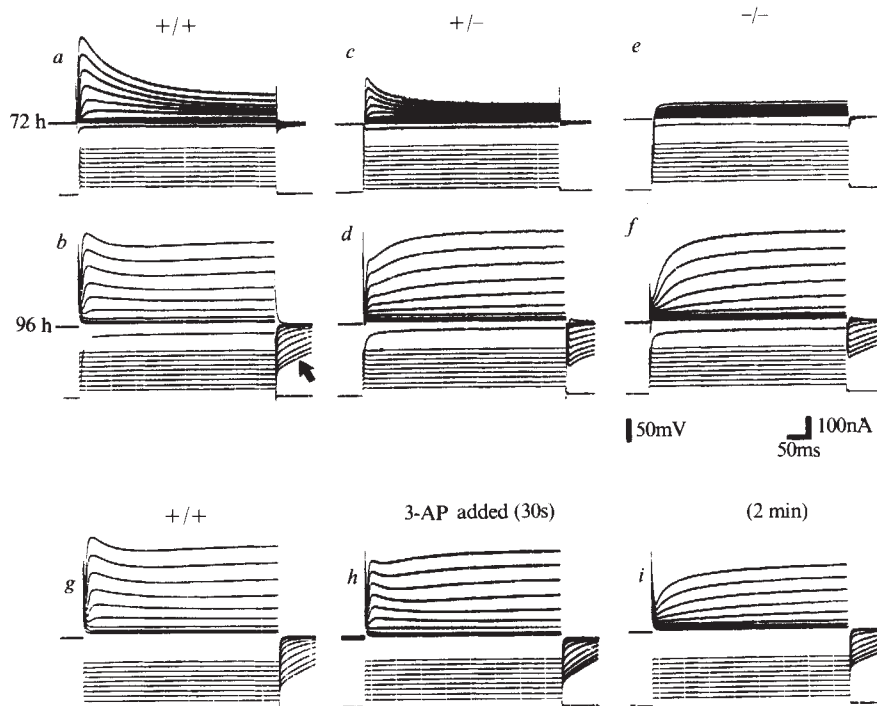
revealed no obvious differences either in activation threshold or channel conductance.

The development of the two distinct sets of voltage-sensitive potassium channels in the DLM was reported previously⁶. The A-current channels normally begin to appear in the DLM 55 h after pupariation. The total amount of A current increases until 72 h of pupal development, when it reaches its mature value. Throughout pupal development the kinetics remain unchanged. Figure 3a shows the wild-type DLM membrane current response to voltage-clamp step pulses in a 72-h cell. The current response consists mostly of A current.

During the next 24 h of development, another set of potassium channels appears in the membrane. These are the channels that carry the slower activating 'delayed rectification' current. Delayed rectification is fully developed at 96 h, just before adult eclosion. The response of the membrane to voltage-clamp step pulses seen in 96-h DLM cells (Fig. 3b) consists of the composite response of two currents, the A current and delayed rectification. Note (Fig. 3b) that the A current predominates at the beginning of the trace, while the slowly activating delayed rectification predominates for the remainder of the trace. Delayed rectification does not inactivate. The channels close only after the membrane voltage is stepped down to the holding potential. The large current tails observed (arrow, Fig. 3b) are due to current still flowing through the closing channels. The rate at which these tails decay reflects the rate of inactivation at -80 mV. This property of non-inactivation at depolarized voltage levels is a characteristic of delayed rectification in other animal systems such as the squid giant axon¹⁰, as well as in the *Drosophila* DLM.

The Sh^{KO120} and Sh^{KS133} mutations modify this normal development of membrane electrical properties during the pupal period by delaying the appearance of the A-current channels. In a 72-h DLM cell, heterozygous for one of these mutations and the wild-type allele, only half the normal A-current response is present (Fig. 3c). The A-current response is similar for heterozygous animals at the 96-h stage; only half the wild-type A-current response is present (Fig. 3d). This reduced

Fig. 3 Comparison of DLM membrane properties during pupal development in wild-type (+/+), heterozygous +/ Sh^{KO120} (+/-), and homozygous mutant Sh^{KO120}/Sh^{KO120} (-/-) pupae. The voltage is shown in the lower traces; membrane current in the upper traces. The holding potential is -80 mV in all cases. *a*, In wild-type pupae the A-current response is mature at 72 h of pupal development; *b*, delayed rectification is mature and is seen added to the A-current response at 96 h of pupal development, just before the eclosion of the adult. The characteristics of delayed rectification are the large, maintained (non-inactivating) outward current during the step pulse, and the large, slow current tails at the end of the step pulse (arrow). *c* and *d*, Membrane current response of heterozygous pupae. At 72 and 96 h only half the A-current response is present relative to wild-type. Delayed rectification is unaffected by the presence of the mutant gene; in *d* at 96 h, the full current response of delayed rectification is present. *e*, *f*, Membrane current response of homozygous mutant pupae. The A current is not present at either 72 or 96 h of pupal development. Comparison of the rising phase of the current traces in *b*, *d* and *f* shows the presence of all, half or none of the A current at 96 h of pupal development. In *f*, at 96 h, the full current response of delayed rectification is present. In *a-f* the current elicited by a hyperpolarizing pulse to -180 mV is shown (below baseline trace) but the voltage record is not shown. *g*, The response of a 96-h wild-type muscle cell before, and *h*, *i*, at various times after the application of 10 mM 3-aminopyridine (3-AP) to the preparation. The A current was partially blocked within 30 s (*h*). By 2 min (*i*) the drug completely blocked the A current but only partially blocked delayed rectification.



A current is almost obscured by the unreduced delayed rectification current. The development of delayed rectification follows a normal time course and is comparable to wild-type at the 96-h stage. Compare the large, maintained outward currents and characteristic current tails at the end of the step pulses in Fig. 3b and d.

In animals homozygous for one of these mutations, the deviation from normal development is most dramatic; the A current is not present in the DLM cells at any stage of pupal development (Fig. 3e,f). Thus, in these animals the membrane response at 72 h of development is almost purely passive; no active channels are present (Fig. 3e). At 96 h, only delayed rectification is present (Fig. 3f). This result clearly shows the independence of the two potassium current-carrying systems. The records shown in Fig. 3 are typical. In wild-type muscle cells the peak of the A current at 96 h was 90–120% of the maximum value of delayed rectification when the membrane was stepped to +20 mV ($N = 9$); in heterozygous cells the A current was 48–75% of delayed rectification ($N = 8$). The A current was totally absent in 19 of 20 homozygous mutant cells in the pupa. (The small current present in one cell had apparently wild-type kinetics.) In contrast to the pupa, the A current is present in most Sh^{KO120}/Sh^{KO120} and Sh^{KS133}/Sh^{KS133} muscle cells after 48 h of adult life. The most direct interpretation of these two mutations is that they delay the developmental appearance of the A-current channels, perhaps by affecting the control of gene expression. However, the possibility cannot be excluded that the mutations affect the biophysical properties of the channels in some, as yet, unknown way.

The separation of the A current from delayed rectification has previously been achieved using pharmacological agents^{13–15}. One such agent, aminopyridine, selectively blocked the A-current response in several different types of animals. A similar effect was observed in wild-type *Drosophila* DLM (Fig. 3g, h, i). The A current was completely blocked 2 min after the application of 3-aminopyridine (10 mM) to 96-h wild-type pupae (Fig. 3i). Delayed rectification, however, was largely unaffected. Note the continued presence of a large maintained outward current, and the inward current tails following cessation of the voltage-clamp step pulse in Fig. 3i. Aminopyridine drugs had previously been used to phenocopy (mimic) the Shaker behavioural phenotype (abnormal twitching) and aberrant neuromuscular physiology (the excessive release of neurotransmitter in response to a nerve impulse)².

Shaker mutations probably affect the A current in neurones as well as muscle cells. The A current is the main repolarizing current in muscle spikes (L.S., unpublished data). Larval motoneurones² and the adult giant interneurone in *Shaker* animals¹⁶ have defective spike repolarization—this can also be phenocopied in wild-type neurones by aminopyridine drugs¹⁶. Thus the various Shaker phenotypes in larvae, pupae and adults are probably all caused by abnormal A currents.

We thank Rick Aldrich and John Thomas for helpful comments on the manuscript, and Charles Stevens for advice and the loan of equipment. This work was supported by USPHS grants NS-14887 and NS-07314 and by a grant from the Muscular Dystrophy Association to L.S.

Received 6 May; accepted 10 July 1981.

- Kelly, L. E. *Nature* **248**, 166–168 (1974).
- Jan, Y. N., Jan, L. Y. & Dennis, M. J. *Proc. R. Soc. B* **198**, 87–108 (1977).
- Wu, C.-F., Ganetzky, B., Jan, L. Y., Jan, Y.-N. & Benzer, S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4047–4051 (1978).
- Wu, C.-F. & Ganetzky, B. *Nature* **286**, 814–816 (1980).
- Salkoff, L. & Wyman, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6216–6220 (1980).
- Salkoff, L. & Wyman, R. *Science* **212**, 461–463 (1981).
- Hagiwara, S., Kusano, K. & Saito, N. *J. Physiol., Lond.* **155**, 470–489 (1961).
- Connor, J. A. & Stevens, C. F. *J. Physiol., Lond.* **213**, 21–30 (1971).
- Neher, E. *J. gen. Physiol.* **58**, 36–53 (1971).
- Hodgkin, A. L. & Huxley, A. F. *J. Physiol., Lond.* **116**, 473–496 (1952).
- Aldrich, R. W., Getting, P. A. & Thompson, S. H. *J. Physiol., Lond.* **291**, 507–530 (1979).
- Pasternak, M. & Kaplan, W. D. *Drosoph. Inf. Serv.* **43**, 62 (1968).
- Thompson, S. H. *J. Physiol., Lond.* **265**, 465–488 (1977).
- Yeh, J. Z., Oxford, G. S., Wu, C. H. & Narahashi, T. *Biophys. J.* **16**, 77–81 (1976).
- Schäuf, C. L., Colton, C. A., Colton, J. S. & Davis, F. A. *J. Pharmac. exp. Ther.* **197**, 414–425 (1976).
- Tanouye, M. A., Ferrus, A. & Fujita, S. C. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

Two variant surface glycoproteins of *Trypanosoma brucei* have a conserved C-terminus

G. Matthysens*, F. Michiels*, R. Hamers*, E. Pays† & M. Steinert†

*Laboratorium Algemene Biologie, Instituut Moleculaire Biologie, Vrije Universiteit Brussel, B-1640 Rhode St Genèse, Belgium

†Département de Biologie Moléculaire, Université Libre de Bruxelles, B-1640 Rhode St Genèse, Belgium

Gene duplication and transposition are known to be involved in the mechanism of antigenic variation in *Trypanosoma brucei*^{1–3}. However, the structure of the antigens in question—the variant surface glycoproteins (VSGs)—is poorly defined. Limited sequencing data show extensive variation between the N-terminal amino acid sequences of different VSGs^{4,5}, although serological studies indicate common antigenic determinants in the C-terminal portion (refs 6, 7 and N. Van Meirvenne, personal communication). More recently, the existence of a C-terminal hydrophobic tail, which is absent from the isolated glycoprotein, was reported⁸. We have now compared the messenger RNA sequences corresponding to the C-terminal 115 amino acids of two serologically different variants belonging to the same serodeme and report a 35% lack of homology (52% of amino acid positions). However, there is a large homologous area of 33 amino acids making up the C-terminus, which includes the 23-amino acid hydrophobic tail. Our data indicate that this extension is a universal feature in *T. brucei* VSGs. In addition, comparison of two isotypes (serologically similar VSGs) belonging to different serodemes showed almost complete nucleotide homology in the coding sequence of the C-terminal 115 amino acids.

We recently described the construction of cDNA clones containing a DNA insert corresponding to the VSG of two variants of a single clone of *T. brucei*, AnTat 1.1 and AnTat 1.8 (ref. 9). Figure 1 shows the restriction map with the regions that were sequenced indicated. Note that the two ends of AnTat 1.8 have an identical restriction map and probably resulted from a simultaneous insertion of identical sequences in opposite orientations⁹. Figure 2 summarizes the sequencing data of the AnTat 1.8 clone. The amino acid sequence was derived by comparing our data with the published nucleotide sequence of VSG 117 (ref. 8) and with the C-terminal protein sequence of four other variants¹⁰; it also turned out to be the only open reading frame.

Comparison of AnTat 1.8 and the 117 variant indicated almost complete base homology. This is not surprising, given that these two variants are serologically indistinguishable¹¹. Furthermore, if we assume that AnTat 1.8 has the same C-terminal aspartic acid residue as its isotypic variant 117 (at codon position 92), then AnTat 1.8 also has a 23-amino acid C-terminal extension. The latter differs by five amino acids between the two variants although its hydrophobicity is conserved. Thus more base substitutions resulting in amino acid change are tolerated in the hydrophobic tail.

Figure 3 compares the AnTat 1.8 sequence with AnTat 1.1, a serologically different variant belonging to the same serodeme¹². We have tried to maximize homology by aligning the two sequences, taking the cysteine residues as invariant, as the number of amino acid residues between two consecutive cysteines is clearly conserved except between positions 77 and 83. Of the 115 amino acid positions, 55 (48%) were identical (65% nucleotide homology). The longest conserved sequence was in the region coding for the hydrophobic tail: although maintaining its hydrophobicity, 7 of 23 residues were different. The hydrophobic extension seems, therefore, to be a universal

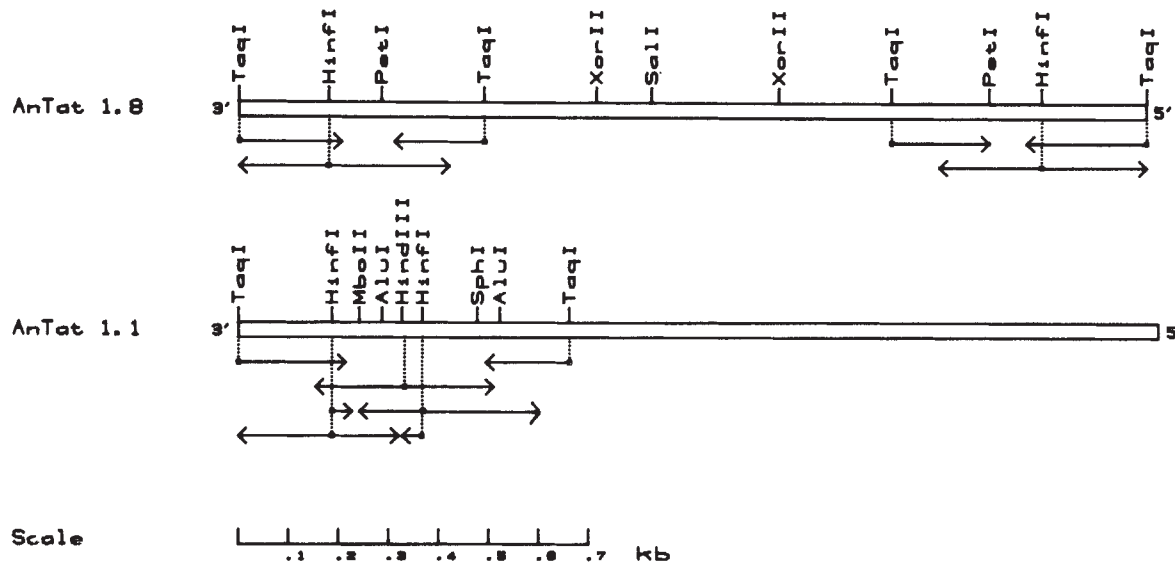


Fig. 1 Sequencing strategy for *T. brucei* cDNA clones of variants AnTat 1.1 (clone A.1 in ref. 9) and AnTat 1.8 (clone A.8a in ref. 9). The direction of each sequencing reaction is shown by horizontal arrows, starting at the corresponding restriction site. The identical restriction map at both ends of the AnTat 1.8 clone probably resulted from a simultaneous insertion of identical sequences in opposite orientations⁹. Each plasmid was initially cleaved with *TaqI*, *HinfI*, or *HindIII*. Restriction sites are from ref. 9 or determined by sequencing. Restriction fragments were end-labelled with ³²P by treatment with bacterial alkaline phosphatase followed by T4 polynucleotide kinase and [γ -³²P]ATP. The double end-labelled fragments were cleaved at an internal restriction site and separated on 6% aqueous acrylamide gels. The eluted fragments were subjected to partial chemical degradation sequence analysis described by Maxam and Gilbert¹⁶ and analysed on 0.3-mm polyacrylamide-urea gels¹⁷.

Fig. 2 Nucleotide and amino acid comparison between two isotopic VSGs, AnTat 1.8 and 117. Dashes indicate codon homology. Codon changes resulting in amino acid substitution are written in full; for neutral base changes, only the base differences are given. Numbers in the figure refer to codon number. The boxed aspartic acid residue has been described^{8,10} as a glycosylated residue known to be at the C-terminus of VSG 117. The so-called 3'-untranslated region located beyond the TAA terminator is written in full. The mRNA poly(A) track is written as (A)n. The AnTat 1.8 sequence is from this work, the 117 sequence from ref. 8, T-cv117.8 cDNA insert. (C)₂₀ indicates the dC tail used to insert the cDNA copy into pBR322.

										10								
AnTat1.8	Aap	Gln	Lys	Gly	Lys	Ser	Pro	Glu	Ser	Glu	Cys	Asn	Lys	Ile	Ser	Glu	Glu	
117	GAT	CAA	AAA	GGC	AAA	TCC	CCT	GAA	AGC	GAG	TGC	AAT	AAA	ATA	TCT	GAG	GAA	
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	

Fig. 3 Nucleotide and predicted amino acid sequence of the *T. brucei* C-terminal end of the AnTat 1.1 and AnTat 1.8 VSGs. To maximize homology, a gap of two codons, indicated by parentheses, was introduced in AnTat 1.1. Solid boxes refer to the cysteine residues, open boxes to the known glycosylated residues. Base homology is indicated by *; other abbreviations as in Fig. 2.

AnTat 1.1 ACA-CCTTTCTTCCCC-TCT----CTTTAAAA-TTTTC-----TTGCTACTTGAAAAAC
AnTat 1.8(1) TTAGCC-----CCCC-TCTTT-CT--AAAAATTTTCCCCCTGCTAAAAATTTTGCTACT-----
AnTat 1.8(n) TTAGCC-----CCCC-TCTTT-CT--AAAAATTTTCCCCCTGCTAAAAATTTTGCTACT-----
117 TT-----TTCCCCCTCTTTTCTT-AAAAATT-C-----TTGCTACTTGAAAA-C

AnTat 1.1 TT-CTGATATATTTTAACACCTTT-----<A> n
AnTat 1.8(1) ---CTGATATATTTTAACACCT-----<A> n
AnTat 1.8(n) ---CTGATATATTTTAACACCT--AAAAAATTCCTCG <A> n
117 -TCCTGATATATTTTAACA <C> 20

feature of *T. brucei* VSGs, allowing at the same time some flexibility in its hydrophobic amino acid content.

The sequences beyond the TAA terminator (the 3'-untranslated region) are shown in Fig. 4. Comparison of these non-coding sequences reveals large differences among different VSGs due to deletions, additions and base substitutions. Examination of the sequences surrounding the putative deletion sites suggests that short direct repeats may be involved in generating these deletions; for example, the AnTat 1.8 sequence contains a direct repeat AAAATTTT which could have been involved in deleting the sequence between these repeats in both AnTat 1.1 and 117 (Fig. 4). Examples of deletions flanked by short direct repeats within noncoding sequences have been reported for mammalian β -like globin genes¹³. That the 3'-untranslated region is a hotspot for recombination is further suggested by analysis of the two ends of the AnTat 1.8 cDNA, which are identical in sequence, except for the 13-nucleotide string AAAAAATTCCTCCG adjacent to the mRNA poly(A) track at the right-hand side of the insert. These sequences may have been generated by linking two double-stranded cDNAs derived from two different mRNAs. There are indications that the VSG mRNAs within a cloned population

consist of molecules of slightly different size (E. Pays, personal communication).

The occurrence of sizeable deletions or insertions during evolution of these genes suggests that the sequences are not essential to VSG mRNA function. In addition, the sequence AAUAAA, which is found in the 3'-untranslated regions of all polyadenylated eukaryotic cellular mRNAs, is absent¹⁴. On the other hand, the sequences TTTT and ATATTT, which occur about 6–10 nucleotides 3' to the poly(A) addition sites in certain eukaryotic genes^{13,15}, are present.

The extensive homology of the C-terminal end of different VSGs, including the hydrophobic tail and 3'-untranslated region, suggests that this region of the VSG is derived from a common ancestral gene. This implies that it could be encoded by exon(s) separated from the rest of the VSG-coding region. Genomic cloning of the respective AnTat 1.1 and AnTat 1.8 genes in both the basic and expression-linked conformation should indicate whether this is the case.

We thank Willy Verheulpen for computer programs and figure layouts, Ignace Lasters for his semiautomated gel reading system, Etienne Wittouck for technical assistance and Milou for typing the manuscript.

Received 21 April; accepted 15 July 1981.

8. Boothroyd, J. C., Cross, G. A. M., Hoeijmakers, J. H. J. & Borst, P. *Nature* **288**, 624–626 (1980).
9. Pays, E. *et al. Nucleic Acids Res.* **8**, 5965–5981 (1980).
10. Holder, A. A. & Cross, G. A. M. *Molec. biochem. Parasit.* **2**, 135–150 (1981).
11. Cross, G. A. M. *Ann. Soc. belge Méd. trop.* **57**, 389–399 (1977).
12. Van Meirvenne, N., Janssens, P. G. & Magnus, E. *Ann. Soc. belge Méd. trop.* **55**, 1–23 (1975).
13. Efstratiadis, A. *et al. Cell* **21**, 653–668 (1980).
14. Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211–214 (1976).
15. Lai, E. C. *et al. Cell* **18**, 829–842 (1979).
16. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
17. Sanger, F. & Coulson, A. R. *FEBS Lett.* **87**, 107–110 (1978).

Changes in the helical repeat of poly(dG-m⁵dC) · poly(dG-m⁵dC) and poly(dG-dC) · poly(dG-dC) associated with the B–Z transition

M. Behe, S. Zimmerman & G. Felsenfeld

Laboratory of Molecular Biology, NIAMDD, NIH, Bethesda, Maryland 20205, USA

Random sequence DNA adsorbed to surfaces such as calcium phosphate has recently been shown to be cut preferentially by nucleases¹ at sites separated by 10.6 ± 0.1 nucleotides², corresponding to the helical periodicity of the DNA. We have now applied a variation of this technique to determine the periodicity of the alternating synthetic polydeoxynucleotides poly(dG-dC) · poly(dG-dC) and poly(dG-m⁵dC) · poly(dG-m⁵dC). We have also obtained X-ray diffraction patterns from fibres of both polymers to compare their helical parameters. Evidence from a variety of physical measurements^{3–5} suggests that in appropriate solvent conditions these polymers can undergo a transition from the usual B form to a new structure tentatively identified with the left-handed Z form. We find by X-ray fibre diffraction that the methylated polymer can adopt a Z form indistinguishable from that of the unmethylated polymer. We further find that when adsorbed to calcium phosphate or calcium oxalate, the B form of each polymer has a repeat of ~ 10.5 base pairs (bp), while the repeat of the Z form is ~ 13.6 bp.

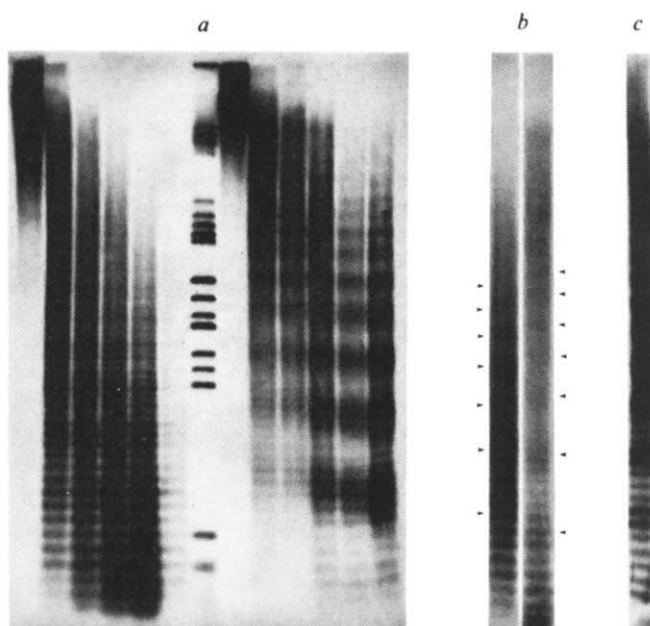
Poly(dG-dC) · poly(dG-dC) undergoes a structural transition in concentrated salt solution³. The transition is accompanied by large changes in circular dichroism³, optical absorbance³ and NMR spectra⁴, as well as in the rate of proton exchange with solvent⁵. Since the recent demonstrations that the oligonucleotides d(CG)₃ and d(CG)₂ crystallize as segments of left-handed helices^{6–8}, the high-salt conformation of poly(dG-dC) · poly(dG-dC) has been thought to correspond to the left-

handed Z form, while the low-salt structure presumably corresponds to the right-handed B form. We have recently examined⁹ the properties of poly(dG-m⁵dC) · poly(dG-m⁵dC), which carries a methyl group in the usual site (CpG) of eukaryotic DNA methylation. The methylated polymer was also found to undergo optical changes consistent with transition to a Z form, but the critical salt concentrations necessary for the transition were much lower than those needed for the unmethylated polymer. Furthermore, the trivalent complex ion, hexamine cobalt, was found to induce the optical transition in both methylated and unmethylated (G · C) polymers in low-salt conditions. Hence, the presumed Z form of both polymers was available for study in moderate conditions of ionic strength, and we have used such conditions in the following studies.

In these experiments, the polynucleotide (uniformly labelled with ³²P during synthesis), is diluted into a solvent which will stabilize either the B or Z form (as judged by circular dichroic spectra) and adsorbed onto crystals of calcium phosphate. The adsorbed nucleic acid is digested with micrococcal nuclease, eluted, glyoxylated and electrophoresed on a polyacrylamide gel.

autoradiograms of the resulting gel for the methylated polymer are shown in Fig. 1a. The preferential cutting of alternate bonds by micrococcal nuclease is even more marked for alternating G-C polymers than for A-T polymers¹⁰; the increment between individual bands corresponds to two nucleotides. The modulation of intensity of larger period reflects the periodicity in accessibility to nuclease and therefore the periodicity of the DNA duplex structure. Tracings of two gel lanes are shown in Fig. 2. It can be seen that between nucleotide positions 0 and 40 there are four major repeats of intensity in the B form pattern, and only three in the Z pattern. The Z repeat thus contains about 4/3 as many nucleotides as the B repeat when adsorbed on calcium phosphate. The peak positions for the Z form are seen (Fig. 1a) to move down on the gel with time of digestion, presumably due to the exonucleolytic action of micrococcal nuclease. Nonetheless, the spacing between peaks, and hence the value obtained for the helical repeat, does not

Fig. 1 a, Autoradiogram of poly(dG-m⁵dC) · poly(dG-m⁵dC) electrophoresed after binding to calcium phosphate and digestion with micrococcal nuclease. Poly(dG-m⁵dC) · poly(dG-m⁵dC) was synthesized from a poly(dI-dC) · poly(dI-dC) template using the large fragment of DNA polymerase I as described previously⁹ with [α -³²P]dGTP included as radioactive label. The product was 300–500 bp long. The nuclease digestion buffer contained 50 mM Tris-HCl, pH 8.0, 10 mM CaCl₂ and 2 mM NaH₂PO₄. For digestion of the B form, ³²P-poly(dG-m⁵dC) · poly(dG-m⁵dC) in 5 mM Tris-HCl, 0.1 mM Na₂EDTA, pH 8.0, was diluted 10-fold with digestion buffer. The presence of Ca²⁺ favours the Z form of the polymer; however, here the B form is rapidly bound to calcium phosphate at low temperature and thus stabilized against the slow conversion to the Z form which would otherwise occur. For digestion of the Z form, the methylated polymer was incubated at 68 °C for 5 min in 5 mM Tris-HCl, 2 mM CaCl₂, pH 8.0, to induce the formation of the Z form and then diluted 10-fold with digestion buffer. The B form was digested at room temperature with 25 U ml⁻¹ and the Z form with 500 U ml⁻¹ of micrococcal nuclease. (The Z form is markedly more resistant to digestion.) The reaction was stopped by addition of EGTA to 20 mM. The samples were then digested with proteinase K and extracted twice with phenol and chloroform. The digested samples were reacted with glyoxal¹⁶. Because of the strong tendency of polymers containing alternating G-C sequences to form self-complementary secondary structures, rigorous denaturing conditions are necessary to obtain good resolution. Although glyoxylation maintains the polynucleotide in the denatured state, it reduces the resolution of the gels compared with what might be expected for homopolymers. The samples were then electrophoresed on an 8% acrylamide, 0.4% bis-acrylamide, 7 M urea gel with 30 mM Tris phosphate, pH 7.2, 1.0 mM EDTA as the gel and electrode buffer. The left lanes are the B form of the methylated polymer digested for 0, 5, 10, 20, 50 and 100 min, the middle lane shows sizing standards, and the right lanes are the Z form digested for 0, 5, 10, 20, 50 and 100 min, respectively. **b**, Autoradiogram of ³²P-poly(dG-dC) · poly(dG-dC) electrophoresed after digestion with micrococcal nuclease while bound to calcium phosphate (B form, left; Z form, right). The polymer was prepared, digested and processed as described above for the methylated polymer except that to induce the Z form of ³²P-poly(dG-dC) · poly(dG-dC) the polymer was incubated at 68 °C for 5 min in 5 mM Tris-HCl, pH 8.0, 20 μ M Co(NH₃)₆Cl₃. The polymer was 300–500 bp long. The digestion buffer in the case of the Z form also contained 0.5 mM Co(NH₃)₆Cl₃. **c**, Autoradiogram of ³²P-poly(dG-dC) · poly(dG-dC) electrophoresed after digestion with DNase II while bound to calcium oxalate. The radioactive polymer was added to 50 mM Na acetate, pH 5.5, 10 mM CaCl₂, 7.5 mM Na₂ oxalate, and digested at room temperature with 0.2 μ g ml⁻¹ of DNase II for 10 min. The reaction was stopped by raising the pH to 8.5 with Tris base and adding 20 mM EGTA to dissolve the calcium oxalate. The digested samples were then processed for electrophoresis as described above. Z form ³²P-poly(dG-dC) · poly(dG-dC) adsorbed to calcium oxalate was not cleaved by DNase II at enzyme concentrations up to 20 μ g ml⁻¹, despite the fact that the enzyme is quite active in digesting DNA in the presence of Co(NH₃)₆Cl₃.



The polymer was prepared, digested and processed as described above for the methylated polymer except that to induce the Z form of ³²P-poly(dG-dC) · poly(dG-dC) the polymer was incubated at 68 °C for 5 min in 5 mM Tris-HCl, pH 8.0, 20 μ M Co(NH₃)₆Cl₃. The polymer was 300–500 bp long. The digestion buffer in the case of the Z form also contained 0.5 mM Co(NH₃)₆Cl₃. **c**, Autoradiogram of ³²P-poly(dG-dC) · poly(dG-dC) electrophoresed after digestion with DNase II while bound to calcium oxalate. The radioactive polymer was added to 50 mM Na acetate, pH 5.5, 10 mM CaCl₂, 7.5 mM Na₂ oxalate, and digested at room temperature with 0.2 μ g ml⁻¹ of DNase II for 10 min. The reaction was stopped by raising the pH to 8.5 with Tris base and adding 20 mM EGTA to dissolve the calcium oxalate. The digested samples were then processed for electrophoresis as described above. Z form ³²P-poly(dG-dC) · poly(dG-dC) adsorbed to calcium oxalate was not cleaved by DNase II at enzyme concentrations up to 20 μ g ml⁻¹, despite the fact that the enzyme is quite active in digesting DNA in the presence of Co(NH₃)₆Cl₃.

change with time. The B form, adsorbed to calcium phosphate, is more resistant to such exonuclease action. Figures 1 and 2 also show that the major periodic variation in intensity is much more marked for the Z form than for the B form, independent of the extent of digestion. Perhaps this indicates that the Z form is accessible to attack over a more limited range of angles, relative to the binding surface, than is the B form. We have found, however, that the amplitude of the periodic variation is a function not only of the DNA conformation, but of the choice both of enzyme (Fig. 1c) and adsorbing surface. Analysis of the amplitude in terms of the conformation of the substrate will therefore require additional data.

The repeat is more accurately measured by direct counting of nucleotide bands between peaks. The value for poly(dG-m⁵dC) · poly(dG-m⁵dC) is 10.4 ± 0.4 nucleotides if the polymer is in the B form when adsorbed. The repeat is 13.6 ± 0.4 if the polymer is adsorbed while in the Z form. The data have also been analysed by autocorrelation analysis or by fitting each of the tracings in Fig. 2 to a single damped sinusoidal curve¹¹, giving numbers falling well within these error estimates. Quite similar values are obtained with poly(dG-m⁵dC) · poly(dG-m⁵dC) adsorbed to a different substrate, calcium oxalate. Digestion with micrococcal nuclease results in gel patterns with a repeat of 10.4 ± 0.4 nucleotides in the B form and 14.0 ± 0.4 in the Z form (data not shown).

We have repeated these studies with unmethylated poly(dG-dC) · poly(dG-dC); the results, with somewhat poorer resolution (Fig. 1b), are quite similar. We find a repeat of 10 ± 1 for the B form and 13 ± 1 for the Z form. Also shown (Fig. 1c) is the result of an experiment in which this polymer was adsorbed to calcium oxalate in the B form and digested with spleen acid DNase (DNase II). Calcium oxalate, rather than calcium phosphate, was used because the latter is soluble in the optimum pH range for the action of DNase II. The accentuation of peaks is more marked in digests with this enzyme, and a more precise estimate of the B repeat (10.5 ± 0.3) is possible. Neither DNase I nor DNase II cuts the polymer effectively if it is adsorbed in the Z form.

The unmethylated polymer, poly(dG-dC) · poly(dG-dC), adopts the Z conformation as judged by comparison of its characteristic X-ray fibre diffraction pattern with that of models based on the Z form oligomers^{12,13}. This fibre diffraction pattern can then be used as a 'fingerprint' to identify the presence of the Z conformation. Using this criterion, the methylated polymer is also capable of entering the Z conformation: the X-ray fibre diffraction patterns of the methylated (Fig. 3a) and unmethylated (Fig. 3b) polymers are virtually identical.

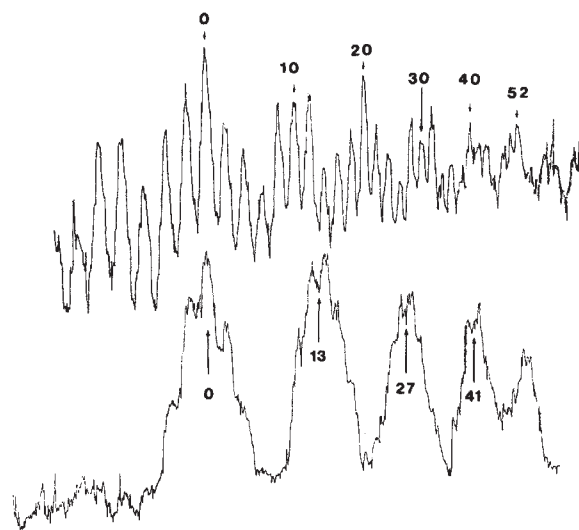


Fig. 2 Densitometer tracing of the autoradiogram in Fig. 1a. Top, B form of poly(dG-m⁵dC) · poly(dG-m⁵dC); bottom, Z form of the methylated polymer. The value of 0 was arbitrarily assigned to the large peak from which we began individual peak to peak counts. The arrows indicate the approximate midpoints of the major cleavage peaks. The tracings are aligned so that bands corresponding to the same number of nucleotides coincide in the top and bottom tracings. Note that bands are separated by 2-nucleotide increments. The positions of the first major peaks fortuitously coincide because of exonuclease action (see text).

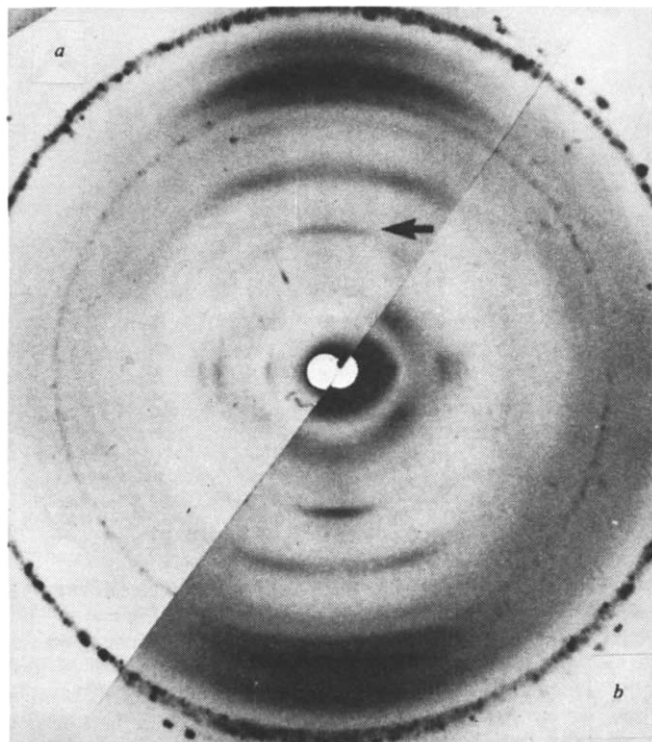


Fig. 3 X-ray fibre diffraction patterns of poly(dG-m⁵dC) · poly(dG-m⁵dC) (a) and poly(dG-dC) · poly(dG-dC) (b) in the Z conformation. The arrow in a indicates the characteristic reflection originating from the dinucleotide repeating unit. This reflection has a Bragg spacing of 7.1 ± 0.1 Å in either a or b, which may be compared with the value of 7.25 Å from Arnott *et al.*¹³ for the unmethylated polymer. Fibres are tilted $\sim 6^\circ$ from the normal to the X-ray beam. The ring of diffracted intensity is from a calcite standard. The fibre in a contains $\sim 4.5\%$ NaCl by weight and is photographed at 52% relative humidity. The fibre in b has no excess salt but is in the sodium form; it is photographed while immersed in 60% v/v ethanol to elicit the Z form. The intensity near the centre of b is due to solvent swelling. The same pattern has been obtained from fibres prepared as in a.

There is a discrepancy between the helical repeat of Z DNA derived from diffraction (12 bp per turn) and the helical periodicity we estimate by nuclease digestion after adsorption to surfaces (13.6–14.0 bp per turn for the methylated polymer, ~ 13 for the unmethylated polymer). It may be recalled that in the case of natural DNA also, the helical repeat based on nuclease digestion of adsorbed material is greater than that determined by fibre diffraction (10.6 compared with 10.0 bp per turn respectively)^{2,14}. Possible effects of close packing in the fibre, or of the perturbation that might accompany adsorption, must be kept in mind. In the case of natural DNA, Rhodes and Klug² observed a cutting repeat for adsorbed material that did not vary with the choice of adsorbing surface, and suggested that the DNA repeat was not altered by binding to the surface. Furthermore, the repeat value obtained agreed well with that found in solution from measurements of the topological winding number of covalently closed DNA¹⁵.

Our data strongly suggest that Z-DNA in solution and in fibre have different helical repeats. Although we find that the repeat of poly(dG-m⁵dC) · poly(dG-m⁵dC) in the presumed Z form is the same whether adsorbed to calcium phosphate or oxalate, it is probably premature to conclude that the repeat of the form reported here is identical to that in solution. Our results and those of others^{12,13} support the conclusion that both methylated and unmethylated (CpG) polymers can assume the left-handed conformation in fibres. Our results also show that the unusual optical changes seen in the transition in solution are accompanied by a large change in helical repeat. This may reflect the presence in solution of a structure related, but not identical to, the fibre form.

M.B. is a Postdoctoral Fellow of the Jane Coffin Childs Fund.

Received 8 May; accepted 31 July 1981.

1. Liu, L. F. & Wang, J. C. *Cell* **15**, 979–984 (1978).
2. Rhodes, D. & Klug, A. *Nature* **286**, 573–578 (1980).
3. Pohl, F. M. & Jovin, T. M. *J. molec. Biol.* **67**, 375–396 (1972).
4. Patel, D. J., Canuel, L. L. & Pohl, F. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2508–2511 (1979).
5. Ramstein, J. & Leng, M. *Nature* **288**, 413–414 (1980).
6. Wang, A. H.-J. *et al. Nature* **282**, 680–686 (1979).
7. Drew, H., Takano, T., Tanaka, S., Itakura, K. & Dickerson, R. E. *Nature* **286**, 567–573 (1980).
8. Crawford, J. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 4016–4020 (1980).
9. Behe, M. & Felsenfeld, G. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1619–1623 (1981).
10. Rhodes, D. & Klug, A. *Nature* **292**, 378–380 (1981).
11. Knott, G. P. *Comp. Prog. Biomed.* **10**, 271–280 (1979).
12. Wang, A. H.-J. *et al. Science* **211**, 171–176 (1981).
13. Arnott, S., Chandrasekaran, R., Birdsall, D. L., Leslie A. G. W. & Ratliff, R. L. *Nature* **283**, 743–745 (1980).
14. Langridge, R., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F. & Hamilton, L. D. *J. molec. Biol.* **2**, 19–37 (1960).
15. Wang, J. C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 200–203 (1979).
16. McMaster, G. K. & Carmichael, G. G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4835–4838 (1977).

Yeast mitochondrial tRNA^{Trp} injected with *E. coli* activating enzyme into *Xenopus* oocytes suppresses UGA termination

Robert P. Martin, Annie-Paule Sibler & Guy Dirheimer

Laboratoire de Biochimie, Institut de Biologie Moléculaire et Cellulaire du CNRS, 15 rue Descartes, and Université Louis Pasteur, 67084 Strasbourg, France

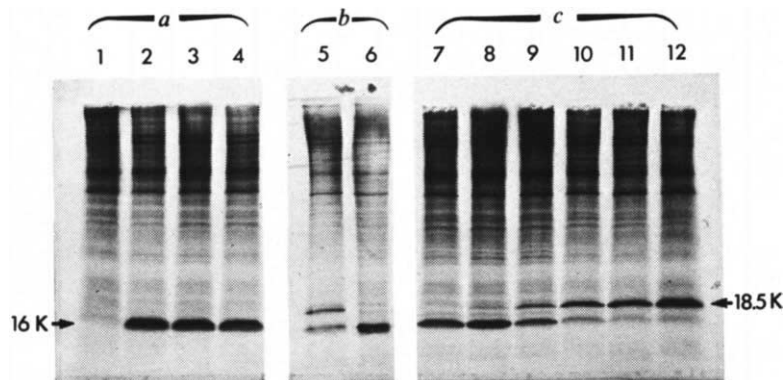
Suzanne de Henau & Henri Grosjean

Laboratoire de Chimie Biologique, Université Libre de Bruxelles, 1640 Rhode-Saint-Genèse, Belgium

It is now clear that the genetic code used in mitochondria differs in a number of ways from the standard code^{1–3}, the most prominent being the use of the opal codon UGA to specify tryptophan^{4–7}. This change in the mitochondrial genetic code is accommodated by the presence of an anticodon U^{CA} in mitochondrial (mt) tRNA^{Trp} from yeast⁸ and *Neurospora crassa*¹. Furthermore, the translation of a mtDNA-encoded mRNA has been achieved in a eukaryotic cell-free system⁹: the mRNA of subunit II of yeast cytochrome *c* oxidase, which contains UGA codons in its reading frame, can be translated into a full-length protein if the system is supplemented with the *Schizosaccharomyces pombe*¹⁰ cytoplasmic UGA suppressor tRNA^{Ser}. It remains to be demonstrated whether a mitochondrial tRNA would be able to function in cytoplasmic protein synthesis. In the case of mt tRNA^{Trp}, this would cause the production of readthrough proteins due to the suppression of the UGA termination codon present in certain cytoplasmic mRNAs. Using the *Xenopus* oocyte microinjection system^{11,12}, we show here that the mt tRNA^{Trp} from *Saccharomyces cerevisiae*, when injected together with rabbit globin mRNA, suppresses UGA termination with high efficiency, thus leading to the production of a β -globin-related readthrough protein of molecular weight (*M*_r) 18,500. However, the suppressor activity of this tRNA within the oocyte cytoplasm is strictly dependent on the co-injection of an exogenous (*Escherichia coli*) acylating enzyme which is needed to charge the mt tRNA^{Trp} *in vivo*. The absence of an endogenous enzyme capable of acylating the yeast mt tRNA^{Trp} suggests that there is a biological barrier for the activity of a mt tRNA in the cytoplasm if a tRNA exchange between the two cellular compartments occurred.

Here we compare the translational effects of cytoplasmic¹³ and mt⁸ tRNA^{Trp} from *S. cerevisiae* and of cytoplasmic opal suppressor tRNA^{Ser} from *S. pombe*¹⁰, after their microinjection into frog oocytes in the presence of rabbit globin mRNAs. The products synthesized in the oocytes were analysed by one-dimensional polyacrylamide gel electrophoresis (Fig. 1). The

Fig. 1 Autoradiogram of ^{35}S -labelled proteins synthesized in *Xenopus laevis* oocytes under the direction of various microinjected macromolecules. Rabbit globin mRNA was prepared according to Marbaix *et al.*²¹. Yeast mt tRNA^{Trp} (anticodon U⁺CA, where U⁺ is a uridine derivative of unknown structure) from *S. cerevisiae* strain IL8-8C was purified as previously described⁸. The *S. cerevisiae* cytoplasmic tRNA^{Trp} (anticodon CmCA)¹³ was purified to homogeneity and the *S. pombe* (strain sup 3-e) opal suppressor tRNA^{Ser} (anticodon U⁺CA, where U⁺ is a mixture of s²U and mcm⁵U)¹⁰ was enriched to 80% purity. Injection of macromolecules into *Xenopus* oocytes and labelling of proteins were performed as described elsewhere^{11,12}. For each sample, 10 oocytes were injected with 0.1 μl per oocyte of saline solution (10 mM Tris-HCl pH 7.8, 1 mM dithioerythritol (DTE), 25% glycerol) containing one or more of the following: rabbit globin mRNA (60 ng μl^{-1}), different tRNAs (in various amounts, as indicated below), purified *E. coli* tryptophanyl-tRNA synthetase (330 ng μl^{-1}). Incubation of the injected oocytes (16–18 h at 19°C) was in the presence of 250 μCi of ^{35}S -methionine (1,340 Ci mmol⁻¹; Amersham) per ml of incubation medium. After extraction of the soluble proteins²², electrophoresis was performed according to Laemmli²³ on 20% polyacrylamide gels. Lane 1, uninjected control oocytes; 2 and 6, oocytes injected with rabbit globin mRNA only. The other lanes in a and b show the products synthesized after co-injection of globin mRNA with yeast cytoplasmic tRNA^{Trp} (200 ng per oocyte, lane 3), yeast mt tRNA^{Trp} (220 ng per oocyte, lane 4), or *S. pombe* UGA suppressor tRNA^{Ser} (120 ng per oocyte, lane 5). In c, globin mRNA and *E. coli* tryptophanyl-tRNA synthetase were injected in the presence of increasing amounts of mt tRNA^{Trp}: 0.11, 0.55, 2.2, 11, 110 and 220 ng per oocyte (lanes 7 to 12, respectively). The fractionation patterns shown in a, b and c originate from different sets of experiments thus the protein bands do not align perfectly. Arrows indicate the 16,000 M_r α - and β -globins (16 K) and the 18,500 M_r β -globin readthrough product (18.5 K).



results indicate that when the uncharged tRNAs are injected, only the suppressor tRNA^{Ser} produces readthrough of the UGA terminator of β -globin mRNA (Fig. 1b, lane 5). Its efficiency of suppression is $70 \pm 5\%$ (ref. 12) when injected at 120 ng per oocyte (that is, 4.8 μM , which corresponds to ~ 3 times the concentration of the endogenous total tRNA in one oocyte¹⁴). In contrast, neither the cytoplasmic tRNA^{Trp}, as expected, nor the mt tRNA^{Trp} (both at 220 ng per oocyte, that is, $\sim 8.8 \mu\text{M}$), induce any detectable β -globin readthrough product (Fig. 1a, lanes 3 and 4, respectively).

In a control experiment, we verified that after microinjection, the yeast cytoplasmic tRNA^{Trp} is aminoacylated by the oocyte enzyme and transfers tryptophan into proteins. This has been shown using methods described for yeast cytoplasmic tRNA^{Phe} (refs 15,16). On the other hand, mt tRNA^{Trp} is not charged at all in the cytoplasm of the oocyte, unless *E. coli* tryptophanyl-tRNA synthetase is co-injected. In experimental conditions where the organelle tRNA^{Trp} is charged, there is a very efficient suppression of the β -globin UGA termination codon. In the presence of 33 ng of tryptophanyl-tRNA synthetase per oocyte (that is, $\sim 0.3 \mu\text{M}$) the amount of β -globin readthrough product synthesized is dependent on the amount of mt tRNA^{Trp} injected (Fig. 1c, lanes 7–12). At 110 ng of mt tRNA^{Trp} per oocyte (lane 11), 93% of the 18,500 M_r protein was observed as opposed to 7% of the 16,000 M_r protein. In fact, this amount can be considered as corresponding to almost 100% readthrough of the β -globin UGA terminator as the remaining 16,000 M_r protein (lanes 11 and 12) corresponds to the translation product of the rabbit α -globin mRNA (termination codon UAA), as revealed by two-dimensional gel electrophoresis (results not shown). The α -globin mRNA is indeed poorly expressed in the oocyte in the absence of added haemin¹⁷.

A quantitative comparison of the readthrough efficiency was made between the mt tRNA^{Trp} and the *S. pombe* opal suppressor tRNA^{Ser} (Fig. 2). Clearly, the mt tRNA^{Trp} is a more efficient suppressor in all ranges of tRNA concentration studied. The largest difference is observed at a low tRNA concentration where as much as 5% of readthrough can be detected with 0.5 ng of mt tRNA^{Trp} per oocyte (that is, $\sim 0.02 \mu\text{M}$) whereas 8–10-fold that amount of the suppressor tRNA^{Ser} is needed to obtain the same effect.

The same is true for the suppression of the UGA stop codon of other mRNAs, as shown by comparing the effects of these two tRNAs on the endogenous translation products of the frog oocyte. Out of the 600 protein spots resolved by two-dimensional gel electrophoresis¹⁸, 59 have been found to be unambiguously affected (that is, appearance of new protein spots or disappearance of pre-existing ones) in the presence of mt tryptophanyl-tRNA^{Trp} (A. P. Sibley, unpublished results), whereas only 16 strong changes have been observed when the opal

suppressor tRNA^{Ser} is injected¹⁹. This difference between the two tRNAs in ability to suppress natural UGA stop codons may result from differences in their intrinsic binding properties due to the nature of the modified U in the first position of their anticodons^{8,10}. Furthermore, other structural differences between the two tRNAs might also be involved in their binding properties and/or in their effect on the ribosome during the translation process. A precedent for this kind of hypothesis is provided by the UGA suppressor tRNA^{Trp} (anticodon CCA) from *E. coli*, in which a sequence alteration in the D-stem confers the ability to translate the UGA nonsense codon in the suppressor strain CAJ64²⁰.

Our results show that yeast mt tRNA^{Trp}, once it has been aminoacylated, is completely functional in all steps of protein synthesis that occur in the cytoplasm of a living cell: for example, codon recognition on the 80S ribosome, interaction with elongation factors, transpeptidation, and competition with the UGA termination factor(s). This poses an important problem if mt tRNA^{Trp} is naturally exported into the cytoplasm of the corresponding eukaryotic cell. In fact, only the charged mt tRNA^{Trp} would be able to interfere in the normal process of cytoplasmic protein synthesis. Apparently, an essential evolutionary barrier which avoids cross-reaction in protein synthesis between the cytoplasm and mitochondria occurs at the level of the charging reaction of the mt tRNA.

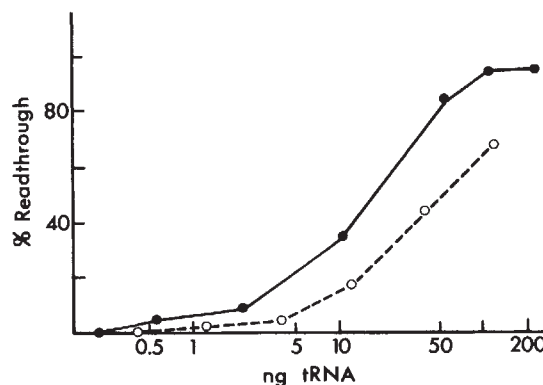


Fig. 2 Percentage of β -globin readthrough protein synthesized in *Xenopus* oocytes as a function of the amount of microinjected mt tRNA^{Trp}. The values (●) were taken from lanes 7 to 12 in Fig. 1c. Each point corresponds to the results obtained for 10 injected oocytes. The ordinate shows effective values of per cent readthrough; the methionine content of the readthrough protein (2 Met residues) and of the β -globin (1 Met) have been taken into account. On the abscissa, the amount of unacylated mt tRNA^{Trp} microinjected per oocyte is given in a logarithmic scale to visualize the minimal tRNA concentration for detectable readthrough. The curve obtained with *S. pombe* UGA suppressor tRNA^{Ser} (○) is given for comparison (results taken from Bienz *et al.*¹²).

The *Xenopus* oocyte microinjection method provides a simple, sensitive and convenient system to assay for the coding properties of a tRNA molecule *in vivo*. Such studies could be used to determine the existence of other coding changes in the mitochondrial genetic code and to evaluate the influence of the reduced number of mt tRNA isoacceptors on the decoding mechanism operating in mitochondria.

We thank Drs G. Marbaix and G. Huez for providing rabbit globin mRNA, Dr G. Keith for yeast cytoplasmic tRNA^{Trp}, Dr J. Kohli for *S. pombe* suppressor tRNA^{Ser}, Drs S. Blanquet and R. Buckingham for *E. coli* tryptophanyl-tRNA synthetase and Professor H. Chantrenne for providing research facilities. This work was supported by grants from the CNRS (ATP 2080 and 4256), from the DGRST (contrat 80.E.0865) and from the Belgian government 'Actions concertées'.

Received 29 May; accepted 23 July 1981.

1. Heckman, J. E., Sarnoff, J., Alzner-Deweerd, B., Yin, S. & RajBhandary, U. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3159–3163 (1980).
2. Barrell, B. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 3164–3166 (1980).
3. Bonitz, S. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 3167–3170 (1980).
4. Barrell, B. G., Bankier, A. T. & Drouin, J. *Nature* **282**, 189–194 (1979).
5. Macino, G., Coruzzi, G., Nobrega, F. G., Li, M. & Tzagoloff, A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3784–3785 (1979).
6. Fox, T. D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6534–6538 (1979).
7. Young, I. G. & Anderson, S. *Gene* **12**, 257–265 (1980).
8. Sibling, A. P., Bordonné, R., Dirheimer, G. & Martin, R. P. *C.r. hebdom. Séanc. Acad. Sci., Paris* **290**, 695–698 (1980).
9. De Ronde, A., Van Loon, A. P. G. M., Grivell, L. A. & Kohli, J. *Nature* **287**, 361–363 (1980).
10. Rafalski, A., Kohli, J., Agris, D. & Söll, D. *Nucleic Acids Res.* **6**, 2683–2695 (1979).
11. Gurdon, J. B. *Nature* **248**, 772–776 (1974).
12. Bienz, M., Kubli, E., Kohli, J., De Henau, S. & Grosjean, H. *Nucleic Acids Res.* **8**, 5169–5178 (1980).
13. Keith, G., Roy, A., Ebel, J. P. & Dirheimer, G. *Biochimie* **54**, 1405–1426 (1972).
14. Brown, D. D. & Littner, E. *J. molec. Biol.* **20**, 95–112 (1966).
15. Gatica, M., Tarrago, A., Allende, C. C. & Allende, J. E. *Nature* **256**, 675–678 (1975).
16. Gatica, M. & Allende, J. E. *Biochem. biophys. Res. Commun.* **79**, 352–356 (1977).
17. Giglioni, B., Gianni, A. M., Comi, P., Ottolenghi, S. & Rungger, D. *Nature new Biol.* **246**, 99–102 (1973).
18. O'Farrell, P. H. *J. biol. Chem.* **250**, 4007–4021 (1975).
19. Bienz, M. *et al. Nucleic Acids Res.* **9**, 3835–3850 (1981).
20. Hirsh, D. & Gold, L. *J. molec. Biol.* **58**, 459–468 (1971).
21. Marbaix, G., Huez, G., Nokin, P. & Cleuter, Y. *FEBS Lett.* **66**, 269–273 (1976).
22. De Robertis, E. M., Partington, G., Longthorne, R. F. & Gurdon, J. B. *J. Embryol. exp. Morph.* **40**, 199–214 (1977).
23. Laemmli, U. K. *Nature* **227**, 681–685 (1970).

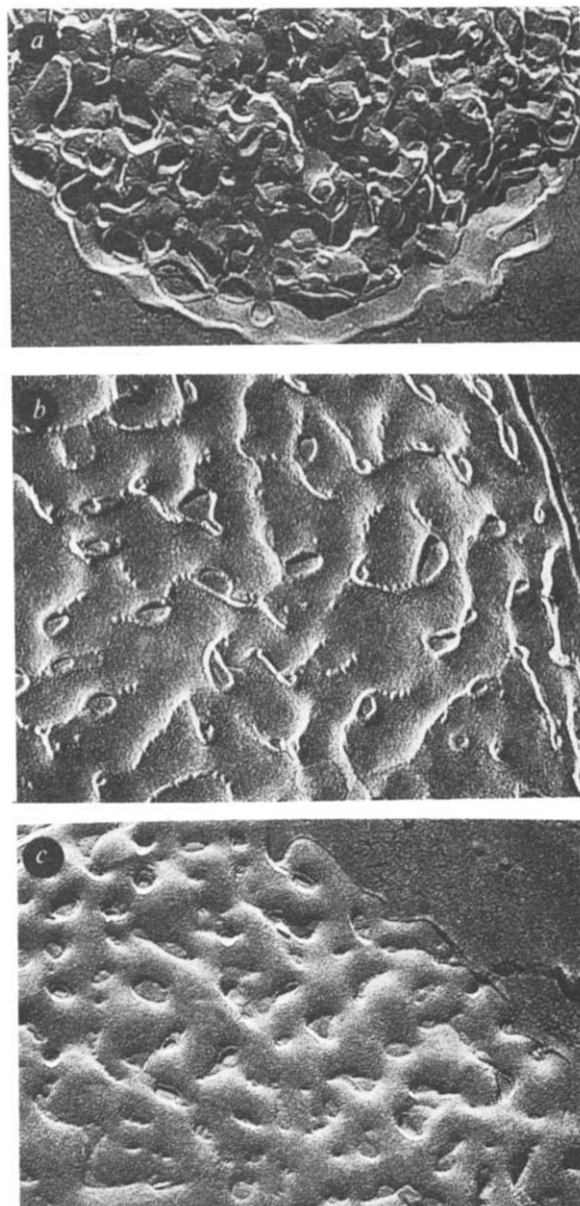


Fig. 1 Freeze-fracture electron micrographs of very rapidly frozen liposomes of bovine cardiolipin and egg phosphatidylcholine (1:1) in 4 mM CaCl₂. Cross-fractures (a) show that these liposomes are full of lipid material ($\times 80,000$). The E face (b; $\times 100,000$) and the P face (c; $\times 100,000$) of the outer bilayer show three types of invagination. Pits and particles or cusps are aligned with both unetched flat areas at the sites of shallow invaginations of the bilayer, and deep invaginations that cross-fracture and etch.

Phospholipid bilayer deformations associated with interbilayer contact and fusion

R. P. Rand*, T. S. Reese† & R. G. Miller‡

* Department of Biological Sciences, Brock University, St Catharines, Ontario, Canada L2S 3A1

† National Institutes of Health, Bethesda, Maryland 20014, USA

‡ Institute of Histology and Embryology, University of Geneva Medical School, 1211 Geneva 4, Switzerland

Bilayers incorporating phospholipids that are capable, in isolation, of forming non-bilayer structures show 'pits and particles' or 'cusps' by freeze-fracture electron microscopy^{1–6}. These points of very high curvature in the bilayer have been interpreted either as inverted micelles within a single bilayer^{2,3} or intersecting bilayers^{4,5}, or as intermembrane sites of attachment^{1,6}. We have used extremely rapid freezing and cleavage at very low temperature to look at the cardiolipin-phosphatidylcholine system and report here a high incidence of cross-fractured liposomes, which shows that even at lower CaCl₂ concentrations they are filled with membranous structures. More interestingly, the arrays of pits and particles are closely aligned and associated with two further kinds of asymmetric bilayer deformation. Flat areas in the bilayer seem to represent sites of particularly adherent interbilayer contact and deep invaginations appear to represent sites where fusion with internal bilayers has occurred.

Liposomes were prepared by combining chloroform solutions of egg phosphatidylcholine (PC) and bovine cardiolipin (CL) to give a molar ratio of 1:1, evaporating the solvent and hydrating the lipid with 65 mM NaCl. The liposomes were bath-sonicated briefly at 0 °C to form a uniform suspension and CaCl₂ added to a final concentration of either 4 or 8 mM; precipitation of the lipid was first seen with 8 mM CaCl₂. The suspension was briefly sonicated and allowed to equilibrate for several hours. A droplet of the suspension was then frozen rapidly without cryoprotectant⁷, cleaved at –150 °C, etched at –105 °C for 5 min with an overlying cold trap, and unidirectionally shadowed with Pt.

Internal and external monolayers of CL/PC liposomes without CaCl₂ are smooth with no apparent substructure of the bilayer. The higher concentration of CaCl₂ gave images consistent with those of Verkleij *et al.*⁵ and with previous X-ray

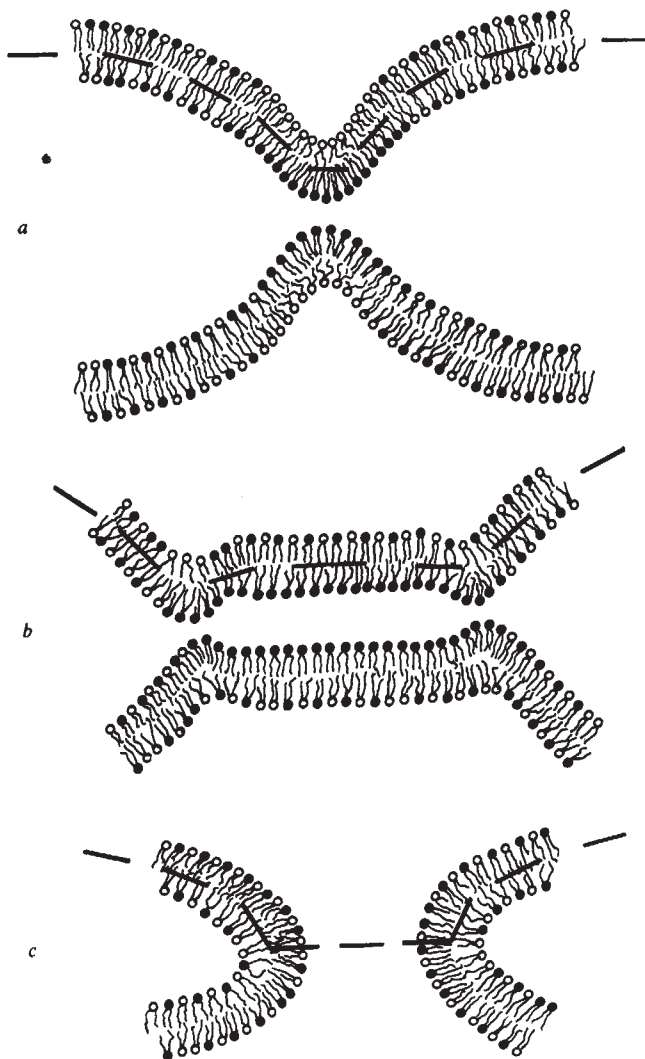


Fig. 2 Our interpretation of the three images of bilayer deformation. *a*, Interbilayer attachment is required to account for the asymmetric punctate deformations whose substructure is unknown. *b*, The flat areas juxtaposed to the curved suggest lateral segregation of the lipid with adherent neutral bilayers in the flat areas and electrostatically repelling bilayers in the curved. *c*, Deep cross-fractured invaginations suggest fusion of outer and inner bilayers.

diffraction experiments which show that this combination of lipids segregates into two distinct phases. In 4 mM CaCl_2 , cross-fractures establish that these liposomes are not single bilayer vesicles but are full of membranous material (Fig. 1*a*). The surfaces of the liposomes show patterns, described earlier¹⁻⁶, formed by linear arrays of projections on shallow ridges in the outer monolayer (E face, Fig. 1*b*) and indentations in valleys on the inner monolayer (P face, Fig. 1*c*) of the outer bilayer of the liposomes. These punctate deformations—'pits and particles'¹ or 'cusps'⁶—are ~ 100 Å in diameter. This study shows that the linear arrays are associated with two further kinds of bilayer deformation; in one of these, the E faces (Fig. 1*a*) show circular ridges of various sizes with steep central holes, resembling volcanoes. The P faces (Fig. 1*c*) show complementarity to the above, with circular depressions of various sizes having steep holes in their bases. The holes are typically surrounded by their adjacent cross-fractured monolayer (appearing as a step around the rim of the holes), and are interpreted as resulting from the etching of aqueous cross-fractured areas. The second type of deformation is similar except that the tops of the volcanoes and the flat-bottomed valleys show no etching and are taken to be

non-aqueous surfaces. The perimeter of both flat areas and etched holes show frequent cusps.

The complete asymmetry of these deformations shows that the outer membrane of the liposome is invaginated. Our interpretation of these invaginations is shown in Fig. 2. The interpretation of the punctate pits and particles and cusps has been recently debated⁶. Both their asymmetry¹⁻⁶ and the fact that they have been seen on liposomes that contain lipid material (refs 1, 2 and Fig. 1*a*) and on bilayers that touch⁶, but have not yet been shown to exist on unilamellar vesicles⁶, strongly support the interpretation that interbilayer contact is required for their formation and stability. Whether that contact represents close apposition of deformed bilayers¹, the formation of partial or complete micelles² or formation of some other structure¹³, or all of these, can be answered only when their substructure has been resolved.

We interpret the flat unetchable surfaces, surrounded by curved bilayers, as areas of stable contact between bilayers or possibly of partial fusion. The fact that they are surrounded by cusps around their perimeter suggests that they may not be single bilayer diaphragms. The juxtaposition of flat and curved bilayers suggests that lateral segregation of the lipid has occurred. Recent measurements of the forces of interaction between lipid bilayers^{8,9} show that the forces themselves can result in lipid segregation. The flat areas suggest that these contain neutral lipid and consequently adhere strongly to apposed bilayers. The curved bilayers could contain charged lipids and experience electrostatic repulsion from internal bilayers^{8,10}.

The etchable holes represent deep invaginations that cross-fracture as the curvature of the bilayer precludes the cleavage plane following the mid-bilayer plane into the liposome. They seem to be sites where the external bilayer has fused with the internal bilayers.

The appearance of two additional kinds of bilayer deformation may be due to the absence of cryoprotectant and the use of very rapid freezing. The close association and alignment of these three kinds of deformation suggest that they represent different stages of bilayer interaction and fusion. We support the idea that the punctate deformations represent the initial critical stage of fusion^{4,5}. Whatever their structure, their curvature is much greater than that seen even in the smallest of bilayer vesicles¹¹. The concave side of such deformations could be packed with phospholipids capable of forming an inverted micelle. Ca-CL could do this in the present system^{5,12}, as could phosphatidylethanolamine (PE) in PC/PE mixtures^{3,6}, where we have shown that bilayer interaction itself could result in the focal accumulation of PE in the area of closest approach of two vesicles⁸. However, the lipid on the convex side is stressed in a direction that would tend to expose the hydrocarbon chains. This could result in the accumulation of lipid in the area that could accommodate such a curvature, as suggested by Fig. 2. Alternatively, close apposition of two such similarly destabilized monolayers could result in the fusion of their hydrocarbon portions and lead, through a series of structural transitions, to the fusion of the bilayers.

We appreciate the help of Mr Bill Graham and Mr Alan Tivey. R.P.R. acknowledges the support of the Natural Sciences and Engineering Research Council of Canada.

Received 26 May; accepted 20 July 1981.

1. Miller, R. G. *Nature* **287**, 166–167 (1980).
2. Verkleij, A. J., Mommers, C., Leunissen-Bijvelt, J. & Ververgaert, P. *Nature* **279**, 162–163 (1979).
3. De Kruijff, B. *Biochim. biophys. Acta* **555**, 200–209 (1979).
4. Verkleij, J., Mommers, C., Gerritsen, W. J., Leunissen-Bijvelt, L. & Cullis, P. P. *Biochim. biophys. Acta* **555**, 358–261 (1979).
5. Verkleij, A. J., van Echteld, C. J. A., Gerritsen, W. J., Cullis, P. P. & de Kruijff, B. *Biochim. biophys. Acta* **600**, 620–624 (1980).
6. Hui, S. W. & Stewart, T. P. *Nature* **290**, 427–428 (1980).
7. Heuser, J. E. et al. *J. Cell Biol.* **81**, 275–300 (1979).
8. Rand, R. P. *A. Rev. Biophys. Bioengng* **10**, 227–314 (1981).
9. Parsegian, V. A., Fuller, N. & Rand, R. P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2750–2754 (1979).
10. Cowley, A. D., Fuller, N. L., Rand, R. P. & Parsegian, V. A. *Biochemistry* **17**, 3163–3168 (1978).
11. Szoka, F. & Papahadjopoulos, D. *A. Rev. Biophys. Bioengng* **9**, 467–508 (1980).
12. Rand, R. P. & Sengupta, S. *Biochim. biophys. Acta* **255**, 484–492 (1972).
13. Hui, S. W., Stewart, T. P., Boni, L. T. & Yeagle, P. L. *Science* **212**, 921–922 (1981).

MATTERS ARISING

V-test is not a statistical test of 'homeward' direction

THE V-test¹ is widely used by behaviourists to evaluate circular data from studies of animal orientation. However, the statistic has been almost universally misinterpreted.

Batschelet¹ very clearly defined the scope of the test when he wrote: "the null hypothesis that we are going to test is randomness, which means that the angles of the sample are independent observations from a uniform circular distribution". However, he further stated that "the V-test leads to significance only when there is sufficient clustering around the predicted direction". It has been this statement, apparently, that has led to some misunderstanding. We can reject the null hypothesis when we have sufficient clustering around the predicted direction. However, the mean direction may still be significantly different from the predicted direction.

Nevertheless, it has been generally assumed that the test will determine whether a sample is oriented in a predicted direction. For example, in four recent papers², the V-test was used, in each case as if rejection of the null hypothesis supported the conclusion that animals were orienting in the appropriate direction. Samples with significant V probabilities were commonly called 'homeward directed'. There has even developed a myth that the V-test complements the Rayleigh test, the latter determining whether a sample is nonrandom, and the former whether the nonrandom sample is properly directed. Thus, Hartwick *et al.*³ said that their data were "evaluated for nonrandomness by the Rayleigh test" and "for homeward directedness by the V-test".

This, unfortunately, is not valid. The null hypothesis of the V-test is the same as for the Rayleigh test, as we have pointed out. The two tests use different methods, but they have the same function. It is true that the calculation of the V statistic uses a predicted direction, but only so that the sample's component on that axis can be compared with the distribution of cosines from "samples from a uniform circular population" (ref. 4). In rejecting the null hypothesis, the researcher can claim that his data are not from a uniform circular population but he cannot claim that the data are from a population of samples oriented in the predicted direction.

A test that addresses the latter question is described elsewhere⁵. The null hypothesis states that the sample mean is drawn from the population of means around the predicted direction. After the nonrandomness of the sample has been demonstrated by either the Rayleigh test or the

V-test, the Watson test can be used to determine whether a confidence interval around the sample mean includes the predicted direction. If the theoretical mean is not within the 95% confidence interval, then the null hypothesis is rejected, and the alternative hypothesis accepted, that the sample mean is not in the predicted direction.

As both tests use the sample mean's component in the predicted direction, sometimes a misinterpreted V-test may seem to give the same results as the Watson test. However, this is not always so; a case in point is the paper by Mather and Baker⁶. Here, the mice transported in the reversed magnetic field had a second-order mean direction of 131°, which is significant by the V-test, using as the predicted direction 180°, the reverse of that expected for the controls. The authors stated that the orientation of the experimental animals was "significantly clustered around 180°". However, use of the more appropriate Watson test shows that the 95% confidence interval around the experimental sample does not include 180°. Consequently, there is no evidence for precisely reverse orientation. From the V-test, we know that the sample mean has a nonrandom component in the predicted direction, but we cannot conclude that the mean itself is in the predicted direction. Thus we cannot conclude anything about the appropriateness of the orientation to the experimental manipulation.

Clearly, Mather and Baker are not alone in their misinterpretation; many others in the field have also used the V-test in the same way. Now seems like a suitable time for reform.

DANIEL J. ANESHANSLEY
T. S. LARKIN

Section of Neurobiology and
Behavior,
School of Electrical Engineering,
Cornell University,
Ithaca, New York 14850, USA

JANICE G. MATHER
R. ROBIN BAKER

Department of Zoology,
University of Manchester,
Manchester M13 9PL, UK

1. Batschelet, E. in *Animal Orientation and Navigation* (eds Galler, S., Schmidt-Koenig, K., Jacobs, G. & Belleville, R.) 63-65 (NASA, Washington DC, 1972).
2. *Animal Migration, Navigation, and Homing* (eds Schmidt-Koenig, K. & Keeton, W.) (Springer, Berlin, 1978).
3. Hartwick *et al.* in *Animal Migration, Navigation, and Homing* (eds Schmidt-Koenig, K. & Keeton, W.) 107 (Springer, Berlin, 1978).
4. Durand, D. & Greenwood, J. J. *Geol.* **66**, 229-238 (1958).
5. Batschelet, E. *Statistical methods for the Analysis of Problems in Animal Orientation and Certain Biological Rhythms*, 29-30 (American Institute for Biological Sciences, Washington DC, 1965).
6. Mather, J. & Baker, R. *Nature* **291**, 152-155 (1981).

mal orientation. It is now apparent that for several years many behaviourists, following Batschelet¹, have misinterpreted the V-test by assuming that because there is a nonrandom vector in a predicted direction, the sample mean is not significantly different from that predicted.

Aneshansley and Larkin have used our paper² on magnetic orientation by mice to illustrate their point. After commenting on our application of the V-test, the authors state that no conclusions can be made about the appropriateness of the orientation to the experimental manipulation. While we accept the comments regarding the V-test, we stress that our main conclusions concerning rodent magnetic orientation are still valid: reversal of the direction of the ambient magnetic field during displacement subsequently influences the animals' ability to determine the direction of home. Most important is the fact that experimental manipulation of the magnetic field results in a significant difference in the directional preferences of control and experimental mice (Watson's $U^2_{n,m}$ non-parametric two-sample test). Moreover, both for controls and experimental animals, the mean orientation has a nonrandom component in the predicted direction (that is, 0° and 180° respectively). Also, the mean for the controls is actually in the predicted direction (Watson's F -test). We can only not say (nor did we in our paper) that the mean for the experimental animals was also in the predicted direction.

We agree with Aneshansley and Larkin, therefore, that we have not presented evidence for precisely reverse orientation but that was never our intention. The main point of our paper was to demonstrate a magnetic sense of direction in a rodent and its involvement in route-based navigation. This still stands.

Department of Zoology,
University of Manchester,
Manchester M13 9PL, UK

1. Batschelet, E. in *Animal Orientation and Navigation* (eds Galler, S., Schmidt-Koenig, K., Jacobs, G. & Belleville, R.) 63-65 (NASA, Washington DC, 1972).
2. Mather, J. & Baker, R. *Nature* **291**, 152-155 (1981).

Molecular packing in collagen

IN RECENT papers, data from chemical cross-linking¹ and electron microscopy² were quoted to challenge aspects of the quasi-hexagonal model³ for the molecular packing in collagen. Any models must, of course, account for X-ray, electron

MATHER AND BAKER REPLY—
Aneshansley and Larkin have raised an important point concerning the interpretation of the V-test, a statistical test that is widely applied in the study of ani-

microscopic and chemical data, but we submit that both of these present challenges fail.

The structure illustrated in our earlier paper³, and reproduced by Bailey *et al.*¹, is, as we pointed out, one of several possible quasi-hexagonal packing schemes, all with the same unit cell and consistent with the X-ray diffraction data^{4,5}. One of these schemes⁵, already preferred in cross-linking arguments, is able to incorporate the 1,5,1,5 cross-links^{1,6} to a greater extent than any other hexagonal model including the 'compressed microfibril' model⁷. The paper by Miller and Tochetti⁸ was in the press and thus not available to Bailey *et al.*¹ at their time of writing and the cross-links used by the former were not those to which Bailey *et al.* drew attention. Contrary to the conclusions of the latter, we have shown that the cross-linking data are consistent with quasi-hexagonal packing, and support the particular packing scheme mentioned above.

The characteristic 3.8-nm row-line in the X-ray pattern from native tendons is often replaced by a row-line in the range 4.5–18 nm in X-ray patterns from fixed or stained specimens^{8–10}. It is possible that the arguments of Squire and Freundlich² are based on such an artefact. It is difficult to preserve the 3.8-nm spacing in specimens prepared for electron microscopy but this has been achieved¹¹ by monitoring the effects of the preparative steps using X-ray diffraction. Furthermore, the observation of Parry and Craig¹² might be due to 4-nm units adhering to the circumference of the fibrils. Finally, Bailey *et al.*¹ refer to a published X-ray pattern¹⁰ of stained tendon showing an 8-nm row-line as strong evidence for an 8-nm lateral unit cell. This is invalid in view of the variability described here.

Hence, while fully accepting that X-ray, electron microscopic and cross-linking data are all relevant, we do not believe that other studies^{1,2} improve on the simple quasi-hexagonal model.

DAVID J. S. HULMES

*Developmental Biology
Laboratory,
Department of Medicine,
Massachusetts General
Hospital,
Harvard Medical School,
Boston, Massachusetts 02114, USA*

ANDREW MILLER

*Laboratory of Molecular
Biophysics,
Department of Zoology,
University of Oxford,
South Parks Road,
Oxford OX1 3PS, UK*

3. Hulmes, D. J. S. & Miller, A. *Nature* **282**, 878–880 (1979).
4. Fraser, R. D. B. & Macrae, T. P. *Katchalsky Symp. Israel* (1980).
5. Miller, A. & Tochetti, D. *Int. J. Biol. Macromolec.* **3**, 9–18 (1981).
6. Davison, P. F. *J. Biol. Chem.* **253**, 5635–5641 (1978).
7. Trus, B. L. & Piez, K. A. *Nature* **286**, 300–301 (1980).
8. Miller, A. & Wray, J. S. *Nature* **230**, 437–439 (1971).
9. Brodsky, B. *et al. Biochim. biophys. Acta* **535**, 25–32 (1978).
10. Nemetschek, Th., Riedl, H. & Jonak, R. *J. molec. Biol.* **133**, 67–83 (1979).
11. Hulmes, D. J. S., Jesior, J.-C., Miller, A. Berthet-Colominas, C. & Wolff, C. *Proc. natn. Acad. Sci. U.S.A.* **78** (in the press).
12. Parry, D. A. D. & Craig, A. S. *Nature* **282**, 213–215 (1979).

SQUIRE AND FREUNDLICH REPLY— Hulmes and Miller have suggested that our observation of a lateral periodicity in electron micrographs of collagen fibrils need not be taken into account in considering possible collagen packing schemes as we are probably studying preparative artefacts. They also imply that the variation of lateral periodicity with different preparative treatments is significant. The following comments may help to put these views into perspective.

(1) Published electron micrographs have frequently shown collagen fibrils which seem to be composed of longitudinal strands^{1,2}. Previous studies of such preparations (for example, by optical diffraction) have failed to demonstrate any regularity between these strands, thus they could have been taken to be artefacts of the preparative procedure in which fibrils which are close-packed *in vivo* are split randomly by the effects of fixation, staining and drying. We have shown using the convolution technique⁶, that such preparations can be very much more regular than has previously been thought. Of course we do not know that collagen molecules occur in such spatially separate bundles *in vivo*; electron microscopy is *per se* a study of preparative artefacts. We would argue, however, that even if molecular clumping in collagen fibrils occurs as a result of the preparative procedure, periodic clumping, however caused, is most likely to be a reflection of a periodic variation in the intermolecular interactions in the native fibril. Thus at the very least our results imply a variation every 50–100 Å in the intermolecular interactions in collagen fibrils. Note that three different preparations gave rise to periodicities in the limited range 50–100 Å.

(2) In the light of the observed variation in the D spacings in our fibril preparations, even within a single fibril, the observed range of values of the lateral spacing was by no means surprising. When analysing our data we considered the postulate³ that collagen fibrils are composed of 80 Å structural units. Despite the observed range of spacings our data are clearly consistent with this postulate. However, the data do not prove this, thus we were careful to note that they are also consistent with a superlattice structure viewed in different directions.

(3) The quasi-hexagonal packing scheme of Hulmes and Miller⁴, which is a very attractive model, is essentially an array of equivalent molecules each ~15 Å in diameter and each making identical interactions with its neighbours. There is no reason *a priori* why in such an array there should be a regular variation in interactions every 50 to 100 Å, nor is there any reason in the quasi-hexagonal model why collagen fibrils should be built up from 40-Å units. Hulmes and Miller suggest this to explain the results of Parry and Craig³ but, although the quasi-hexagonal model has a unit cell with a principal interplanar spacing of 37.8 Å, this is not a structural unit. The largest structural unit in the model, as proposed⁴, is an individual collagen molecule 15 Å in diameter.

(4) We consider it likely that, superimposed on the basic quasi-hexagonal lattice, there is a pattern of intermolecular interactions, including specific cross-links, which defines molecular groupings on a larger scale (that is 50–100 Å). The models of Trus and Piez⁵ represent examples of this kind of structure; indeed Hulmes and Miller⁴ themselves acknowledge such a possibility. We hope that our studies of cryo-sectioned collagen fibrils⁶ will help to define the size and nature of these larger molecular groupings for which X-ray diffraction has so far failed to provide tangible data.

JOHN M. SQUIRE

ALAN FREUNDLICH

*Biopolymer Group,
Department of Metallurgy
and Materials Science,
Imperial College,
London SW7 2BP, UK*

1. Grant, R. A., Cox, R. W. & Horne, R. W. *J. R. microsc. Soc.* **87**, 143–155 (1967).
2. Smith, J. W. & Frame, J. *J. Cell Sci.* **4**, 421–436 (1969).
3. Parry, D. A. D. & Craig, A. S. *Nature* **282**, 213–215 (1979).
4. Hulmes, D. J. S. & Miller, A. *Nature* **282**, 878–880 (1979).
5. Trus, B. L. & Piez, K. A. *Nature* **286**, 300–301 (1980).
6. Squire, J. M. & Freundlich, A. *Nature* **288**, 410–413 (1980).

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 300 words and the briefest of replies, to the effect that a point is taken, should be considered.

1. Bailey, A. J., Light, N. D. & Atkins, E. D. T. *Nature* **288**, 408–410 (1980).
2. Squire, J. M. & Freundlich, A. *Nature* **288**, 410–413 (1980).

BOOK REVIEWS

Means of getting more from less

Duncan Davies

IT IS not difficult to understand the rapid growth of interest in science, technology, R&D, innovation, enterprise and associated risk investment. The citizen in the affluent developed countries has been taught to expect increasing consumption of goods and services. For the past decade, these expectations have been disappointed, in part because the primary producing countries — and especially the oil producers — have used their bargaining power for the first time, and removed a significant slice of the affluent cake. This has helped the money supply to grow faster than the goods supply. It has also reinforced the processes whereby the rate of growth in purchases of goods has fallen below the rate of growth of manpower productivity, with resulting unemployment and general malaise.

The traditional and more obvious remedies are not available. After 40 years of successful demand management (from 1933) there is no longer the headroom for acceptable trading of more inflation for less unemployment. So governments have turned to supply management. Many mineral resources — and especially oil — are in the hands of non-OECD countries, so management here is difficult. Management of manpower resources is inhibited by the powerful veto of many small groups who stand astride long production chains. All that is left is innovation — more of the expected consumption from scarcer resources. So spotlights and expectant eyes are turned on to the innovators. How do they work? Who can fund their work and its exploitation? Where does science come in? Are we educating the right sort of people in the right sort of way? Perhaps, most important of all, can the free market provide an adequate framework, or is voluntarily invested cash too impatient to provide for the bigger entry fees and the longer lead times, now extending up to a century ahead in extreme cases such as fusion power?

The relevant literature can be written in many ways, none of them entirely satisfactory. One can present the viewpoint of an individual, and anecdotally describe how he was educated in line with the friendly and optimistic paradigms of an earlier age, only to emerge into a world that did not fit and was harsher. This is the standpoint of *Public Science — Private View*, an agreeable book by David Budworth. One can write of the adaptive struggles of a nation state to manage its science and their public practitioners and regulators. This is the theme of Philip Gummert's *Scientists in Whitehall*, a

Public Science — Private View. By D.W. Budworth. Pp.183. Hbk ISBN 0-85274-449-8; pbk ISBN 0-85274-452-8. (Adam Hilger, Bristol/Heyden, Philadelphia: 1981.) Hbk £9.95, \$24.75; pbk £5.95, \$14.75. *Scientists in Whitehall*. By Philip Gummert. Pp.245. ISBN 0-7190-0791-7. (Manchester University Press/Humanities Press: 1980.) £14.50, \$38. *Industrial Innovation and Public Policy: Preparing for the 1980s and 1990s*. By Roy Rothwell and Walter Zegveld. Pp.251. ISBN 0-903804-92-1. (Frances Pinter, London: 1981.) £13.50.

scholarly and clear account of a pragmatic, muddled and slow progress. One can look at the problem internationally, and thus try to cover comparative advantage, trade, competition and very diverse patterns of behaviour. This is the much more difficult task attempted by Roy Rothwell and Walter Zegveld in *Industrial Innovation and Public Policy*. Not surprisingly, their book is more of a patchwork of very ill-assorted textiles, arranged with a care that cannot suffice to create a satisfactory pattern.

All three books end with a doubtful and hesitant hope that they have helped others to understand and cope with the problems. To understand, yes; to cope, so far no. Time is one difficulty: substitution even of a simple kind (synthetic fibres, water-bound paints, automobiles for horses) typically takes 50 years or so, whereas taxpayers and voters expect results in one or two years. Tradition is another: monastically based education set up so as to man the British Empire may have difficulty in generating world-competitive mass production; others, who last century lacked either cheap materials (Germany) or plentiful manpower (America) or national victories (France), may have laid foundations more suitable for late-twentieth-century competitiveness, in various ways. Difficulty in assessing good innovative practices also arises from the large number of other good practices equally necessary for their success: engineering, banking, production, marketing, and the motivation, training and deployment of people. So one can compare the costs of the specific inputs to science, technology and innovation, but there is no good measure of the separate value of the outputs, and no shortage of other claimants to responsibility for the final profit or benefit. Patents measure the determination to patent as much as the quality and utility of the material patented. In some respects, then, a conspectus of anecdotal material may, maddeningly to numerate tech-

nologists, be more informative than numbers and graphs.

It is therefore legitimate to take David Budworth's *Odyssey* seriously. He travelled hopefully and enjoyably through scientific education into technological research to the post-doctoral stage, and then became a university teacher. He nearly stayed in America, and wishes that he had, because there technology is valued more highly than in the UK: teamwork and application in industry are thought of as intellectually challenging and socially valid. He moved into a brave but belated attempt to upgrade the tableware industry systematically, and has harsh things to say about Pharisaic comments from government laboratories and university lack of interest. In Shirley Williams's words, the party was over, and he was squeezed out. But his observations are mainly about the research system, the universities and the impact of Rothschild on science. Having seen that the science system is not very responsive in the UK to world market factors, he does not go on to enquire into the question whether the market alone will back the necessary technology. The French and Japanese reckon that it will not; its imperatives are crucial, but insufficient. Accordingly, Budworth has yet to write the *Iliad*. His style and turn of phrase make me hope that he will. So far, so good.

Philip Gummert's book is a valuable quarry, and brings together a lot of useful history. But it makes hardly any reference to systems outside the UK or to the need and basis for selectivity. So, bearing in mind the UK's difficulties, it cannot say very much about where to go next. It is about means, not ends.

In his closing pages, Gummert almost escapes from the parochialism of science policy. He observes (p.228) that ACSP's (Advisory Council for Scientific Policy) 1954 warning of the dangers of turning out too many narrow science specialists was not followed by marked changes in the educational system, which therefore "may still impede the better use of scientists in Whitehall". He continues:

to wonder whether it would not be in the best interests of British Science and British government to broaden the base of experience which is being tapped; interestingly, this is not a question that has attracted much attention among scientists.

This question is central, penetrating and pertinent, and is firmly addressed in the recent Holdgate report on the Scientific Civil Service. In other countries, there is an overlapping and continuous spectrum of people with broad capabilities ranging

from the non-technological cleric, lawyer and financier at one end to the highly-skilled biotechnologist or number theorist at the other. The private sector in the UK has responded similarly, as far as it can in the teeth of our educational specialist conservatism. But in the public sector we have failed to see that technological generalists — the middle members of the spectrum — are the essential allies of the specialists, with whom they are symbiotic and mutually dependent. They are not enemies from an alternative and hostile culture, but they tend to be treated as such because they question the sufficiency (though not the necessity) of specialist freedom. So it is hard to accept Gummert's tentative suggestion that the day may be saved by the Advisory Board for the Research Councils — a wise and distinguished body, whose main job is defined by its title. It is seeking to ally itself more closely with its sister Advisory Council for Applied Research and Development and to use science more effectively (as well as protecting its independence). But the difficulty is that ACARD has no money and hence only influence, not authority. Under the British system, so ably described by Gummert, an ACARD with a lot of financial clout would be a surrogate micro-Cabinet. For better or worse, it would be opposed by Departments and Ministers, whose functions it would duplicate. Accordingly, the technological generalists, who have now begun to

become significant forces (more warmly welcomed by the arts administrators than the traditional scientists), must operate on an informal, club basis when they seek to co-ordinate the work of different ministries. They have had to come in from outside, but before long the civil service itself will produce some top technological generalists of its own. Come back Fulton, all is forgiven, says Holdgate; Whitehall is following. Will the university scientists and their learned societies follow? I think that they will, but it has taken a long time.

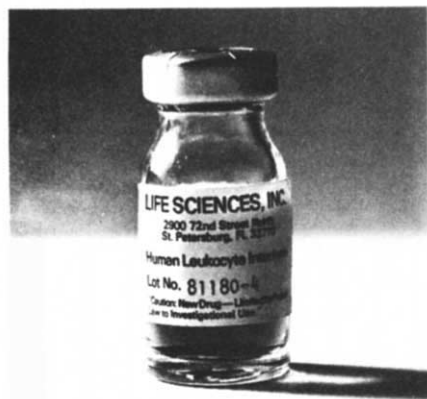
Rothwell and Zegveld are academic technological generalists who will be read by an increasing number of people in Whitehall as well as those working in industry. How far have they provided the intellectual infrastructure that these proud pragmatists can use? A little way, but not far enough. In particular, they are not generous enough in recognizing the crucial role of the private sector — insufficient certainly, on its own, and dependent on government actions. These include the climate provided by the patent system and limited liability, and backing for technical infrastructure with lead times too long for firms living in a world of 8–20 per cent interest rates. But the successful selective systems — France, Sweden, Japan — allow the choice of products and processes to be made by firms, and allow competition to winnow the contenders. This is an area where governments alone have a poor track record — except in wartime, a factor which causes confusion. Academics and public servants tend to underestimate the very basic differences between rules for the military and peaceful application of technology. This useful and valuable book is therefore flawed; it does not merely underestimate the private role, it says

(p.238) “technological laissez faire is presently as much a thing of the past as economic laissez faire”. “Is”, is manifestly wrong; “might be”, “ought to be”, or “will be” would be debatable. Most OECD governments have been and are trying to liberalize technology policies in various ways, and it is at least arguable that the enormous capital requirements for the development of Colorado shale have been made available more effectively through price de-regulation than through President Carter's tax-based Synfuels Corporation. After all, Carter himself tried hard to de-regulate, and failed for lack of political authority. There are many further examples.

But perhaps such cavilling is unworthy. Rothwell and Zegveld lay out all the main issues, and assemble an impressive array of published data and ideas. It is possible to add material about the key role of the international companies in innovation and technology transfer, and then use the structure of their last chapter to consider the public-private sector synthesis so as to look forward and decide on new wayleaves for enterprise and venture capital, alongside government investment of taxpayers' money in basic technology. The application of science and technology requires the individualism of the discoverer, the inventor and the entrepreneur to get the show on the road, and the collective international strength of firms and governments for it to grow and flourish. The unification of these psychologically very diverse capabilities is the key role of the technological generalist. □

Duncan Davies is Chief Engineer and Scientist at the government Department of Industry, London.


**LIFE SCIENCES,
INC.**



Supplying Human
Leukocyte Interferon for
clinical tests and
research.

2900-72nd Street North
St. Petersburg, Florida 33710
(813) 345-9371

Circle No. 15 on Reader Enquiry Card.

Understanding nuclear collisions

D.M. Brink

Nuclear Reactions with Heavy Ions. By Reiner Bass. Pp.410. ISBN 3-540-09611-6/0-387-09611-6. (Springer-Verlag: 1980.) DM 88, \$ 52.

THE study of nuclear reactions initiated by heavy ions has become one of the main areas of activity in nuclear structure physics. The projectile in a heavy ion reaction is a heavy nucleus accelerated as an ion in a cyclotron or some other kind of accelerator, the object being to study interactions between the projectile and a nucleus of an atom in the target. The interaction of α -particles with nuclei is known to exhibit features characteristic of much heavier and more complicated projectiles — the α -particle has sometimes been referred to as the “lightest heavy ion” — and helium beams of high quality and intensity have been available for many

years. Now there are a number of different facilities available around the world for accelerating heavy nuclei.

Professor Bass's book gives an excellent survey of the present state of our understanding of heavy ion reactions. He covers most aspects of the subject, but excludes reactions initiated by relativistic heavy ions as well as more specialized theoretical topics.

The de Broglie wavelength associated with the relative motion of two nuclei in a heavy ion collision is much smaller than the radii of the nuclei involved and their orbital angular momentum is large. In these circumstances it is possible to use semi-classical or even classical mechanics to describe a collision. The wave effects associated with quantum theory are noticeable only in certain kinds of experiments. The first chapter of the book

gives a summary of the results of both classical and quantum scattering theory needed for subsequent discussions. In addition, some experiments involving collisions between lighter nuclei can be understood by supposing that there is an intermediate state in which the nuclei stick together forming a kind of nuclear molecule, and the author presents evidence for this quasi-molecular picture.

About one-half of the book is concerned with quasi-elastic scattering, that is the simpler kinds of scattering processes where the interacting nuclei are left in their ground state after the collision, or in a low excited state. Also included in this category are reactions where one or several nucleons are exchanged between the nuclei during the collision. A great deal of experimental and theoretical effort has been put into studies of quasi-elastic scattering and Professor Bass has collected together the most interesting results. There are a number of open questions here and, as well as reviewing the present state of understanding, the author provides a very complete set of references.

Deep inelastic scattering and fusion are other processes which can occur in a heavy ion collision, and Professor Bass gives a clear review of current experimental results and theoretical ideas. In deep inelastic scattering two fragments similar in mass to the initial nuclei emerge from the collision, but almost all the energy of relative motion is transferred into internal excitation. This transfer takes place in a very short time, comparable with the transit time of a nucleon across a nucleus. There has been intense discussion about the mechanism of this energy loss: is it due to collisions between nuclei or does it involve the interaction of nucleons in one nucleus with the average potential field due to the other nucleus?

The book also has an interesting chapter on superheavy nuclei. For a number of years there has been speculation that nuclei with atomic number Z in the range 112–126 might exist and be moderately stable. Such nuclei could be formed in a fusion reaction between two lighter nuclei. Present studies of fusion reactions have shown that it is not easy to get two nuclei to fuse permanently into a compound system. Usually they seem to fuse temporarily and then separate again, but with a better understanding of the fusion process it might be possible to form superheavy nuclei.

Professor Bass has succeeded in giving a most valuable overview of all these topics. The author's own research is mainly in the fields of deep inelastic scattering and fusion, and these topics are especially well discussed and a balanced view is presented. The references are complete throughout. It is a well-planned book which will be very useful to research workers in the field.

D.M. Brink is in the Department of Theoretical Physics at the University of Oxford.

Enzyme imbalance

M.C. Scrutton

Enzyme Regulation and Metabolic Diseases. By Francesco Belfiore. Pp.877. ISBN 3-8055-0005-X. (Karger, Basel: 1981.) SwFr. 249, \$149.25.

THE relationship between enzyme deficiency and metabolic disease has been thoroughly explored since Garrod formulated the concept of inborn errors of metabolism more than 70 years ago. In contrast, the possible role of altered enzyme regulation has received much less attention, although the concept is attractive especially for more common disorders having a prolonged time course such as obesity, atherosclerosis and the secondary complications of diabetes. Galton, in particular, has proposed that a defect in the regulatory site of a key enzyme might be important to the development of certain of these conditions and has provided some supporting evidence. An alternative view, and that which is espoused here by Dr Belfiore, is that altered enzyme regulation is "a change in regulatory factor(s) in the presence of normal enzyme(s)". Having thus set the context in which the topic is to be considered, the various mechanisms which might be important to altered regulation are very briefly considered. There then follow four massive sections which examine the evidence for, and possible role of, altered enzyme regulation in obesity, diabetes, the hyperlipoproteinemias, and atherosclerosis and diabetic microangiopathy.

My overall impression is that this book lacks balance and critical insight in the treatment of a concept which is not widely accepted and for which, except in a few instances, there is little clear-cut evidence at the biochemical level. Although widespread changes occur in the levels of both metabolites and enzymes in all the conditions discussed, the most pertinent question is which, if any, of these changes are causal and to what phase of the overall disorder do such changes relate. These questions are extremely difficult to answer but Belfiore's case is not helped by failing to formulate them clearly. The lack of balance is shown clearly by the relative weight given to discussion of general mechanisms, which is minimal, versus the properties, and changes in the level, of specific enzymes in the various disease states examined. It is also exemplified by the failure to give reasonable attention to both of the hypotheses which might explain altered enzyme regulation in the disease states considered, given the absence of compelling evidence favouring either one of them. Since no coherent postulates are presented, the sections relating to specific disease states have become a compendium of information which might conceivably (or even inconceivably!) be relevant. The result resembles that of an old-style Annual Review in which the objective was to detail

SCIENTIFIC BOOKSHOP

H.K. LEWIS can supply works in all branches of Pure and Applied Science. Catalogues on request. Please state interests.

SCIENTIFIC LENDING LIBRARY

Annual Subscription from £8.50
(Available in U.K. only)

Reduced rates for multiple subscriptions.

Prospectus post free on request.

Quarterly List of New Books and new editions added to the Library sent post free to subscribers regularly.

**H.K. LEWIS
& Co. Ltd.**

136 GOWER STREET,
LONDON, WC1E 6BS

Telephone: 01-387 4282
Telegrams: "Publicavit,
London, WC1E 6BS."

Circle No.22 on Reader Enquiry Card.

CAN THE ANALYSIS OF
NOISE PRESENT IN
POWER REACTORS BE
USED AS A TOOL FOR
REACTOR MONITORING
AND DIAGNOSTICS?

READ:
**POWER REACTOR
NOISE**

by
Joseph A. Thie

224 pages • \$38.00 • 6 x 9 Hardbound

**JUST PUBLISHED by the American
Nuclear Society**

Send order form to:

AMERICAN NUCLEAR SOCIETY
555 North Kensington Avenue
La Grange Park, IL 60525 USA

Please take my order for ___ book(s).

☐ Payment in full enclosed.
☐ Bill me and I accept postage and handling charges.

Name: _____

Organization: _____

Street: _____

City: _____

State: _____ Zip: _____

7/81

Circle No.04 on Reader Enquiry Card.

all information published on the topic in the past year. In this case, however, the period covered is 1950–1977 giving rise to nearly 4,500 references in a section which occupies nearly a third of the text.

The collection of a vast amount of data is, then, the author's major contribution. If this book has a use, it is as a catalogue and source of information for those who have the patience to sift the wheat from the chaff. □

M.C. Scrutton is Professor of Biochemistry at King's College, University of London.

A taste of modelling

William L. Jolly

Molecular Shapes: Theoretical Models of Inorganic Stereochemistry. By Jeremy K. Burdett. Pp.287. ISBN 0-471-07860-3. (Wiley: 1981.) £24.05, \$43.25.

ONE of the main efforts in modern inorganic chemistry is the search for simple yet accurate methods for predicting the shapes of molecules. The problem is critical because of the explosion in the number of new structural types that have been synthesized in recent years. In this book Jeremy Burdett briefly describes the more useful of these predictive methods, including the electron pair repulsion approach, the Walsh diagram approach, the Jahn–Teller theorem, the angular overlap method, the crystal field model, the 18-electron rule, the cluster electron counting rules and the fragment formalism. The emphasis is on simple molecular orbital theory, with frequent use of the angular overlap model.

Each topic is discussed in a clear and critical manner, and I am sure that most readers will gain new and useful insights from a reading of any one of the 15 chapters in the book. My only possible complaint is that the book is rather short. Thus Wade's rules for electron-counting in cage compounds are covered in only 13 pages, and the fragment formalism in solid state chemistry (a topic developed by Burdett himself) is discussed in fewer than five. Yet in another sense the coverage of each topic is ideal: the material is presented so interestingly that the reader's appetite is whetted so that he is eager to go to the original literature and learn the details which have been omitted from this book.

To a certain extent, Burdett's approach is idiosyncratic because of its stress on the angular overlap method and arguments *à la* Roald Hoffmann. However that emphasis adds a certain flavour of eccentricity, and I heartily recommend the book to all inorganic and organometallic chemists.

William L. Jolly is a Professor in the Department of Chemistry at the University of California, Berkeley.

Behavioural ecology as competition

T.R. Birkhead

Behavioral Mechanisms in Ecology. By Douglass H. Morse. Pp.383. ISBN 0-674-06460-7. (Harvard University Press: 1981.) \$32.50, £17.50.

THERE are now a number of behavioural ecology books available, but this one differs from most others in being primarily concerned with how animals behave rather than with selfish genes and why particular behaviour patterns have evolved. The main theme is competition for resources, mainly food, and Morse starts by discussing foraging efficiency in some detail. The rest of the book is concerned with the constraints on foraging efficiency, such as the effect of predators, high and low temperatures, reproductive activities and so on.

Chapters 2 and 3, on foraging, are perhaps the best. The main aspects of optimal foraging are clearly stated; some of the predictions of optimal foraging theory are presented and the field and laboratory evidence is objectively examined and assessed. Morse concludes that optimal foraging models based simply on energy maximization are not especially useful in predicting foraging behaviour in the wild because they ignore nutritional factors, competition and predation. While these were pertinent criticisms of optimal foraging studies at the time Morse was writing this book, many of his suggestions have already been followed through. For example, in the past few years both nutrient constraints and the influence of predation have been incorporated into optimal foraging models with some success. In discussing habitat selection (Chapter 4) the similarities between optimal foraging (in particular food and patch choice) and habitat selection are outlined, and the importance of habitat selection in terms of food location is emphasized.

The author next looks at the ways animals avoid becoming food items themselves and considers antipredator behaviours, such as alarm calling, from altruistic and selfish standpoints, and then goes on to review behavioural thermoregulation. Under reproduction (Chapter 7) the author discusses parental care and communal breeding. I was surprised to find no mention of evolutionary stable strategies in the discussion of parental care, or indeed elsewhere, and disappointed that there is no mention of communal breeding in groups other than birds. In the next chapter, on the mechanisms associated

with competition for mates, Morse considers the main types of competitive interactions, visual, auditory and olfactory signals used in such interactions, and finishes by considering sexual size dimorphisms. He suggests that these have evolved as a consequence of sexual selection rather than as a result of selection for diversification in foraging between the sexes. This conclusion is based solely on the fact that relatively few sexual feeding morphs exist. While this is entirely reasonable, it is unfortunate that the author misses this opportunity to point out that there have been marked advances in the comparative approach over the past few years and that there now exist quantitative techniques for disentangling the effects of confounding variables.

Territoriality (Chapter 9) covers some standard topics. The main emphasis here is on feeding territories, and as elsewhere in the book the majority of examples are drawn from studies of birds. Chapter 10 looks at dominance hierarchies in relation to food acquisition, and 11 covers inter-specific competition, namely, the form of interactions, spatial relationships between competing species, changes in niche breadth and community structure. The benefits (but not the costs) of living groups are discussed in the penultimate chapter on feeding advantages and antipredator advantages. Finally, Morse considers the future direction of behavioural ecology, urging field workers to examine differences between populations of the same species, and between individuals, to test some of the ideas of R.L. Trivers and others. Unfortunately this chapter has lost some of its impact in the delay between writing and publication. The same is true for much of the rest of the book; there are few references after 1978 and none after 1979.

Behavioural ecologists thinking about using this book as a course text may find the approach rather narrow and the treatment of some topics somewhat incomplete. Another shortcoming, I feel, is the lack of any attempt (except in Chapter 3) to outline theoretical models or evolutionary concepts such as optimization or ESSs. While few would dispute the importance of food acquisition, the way this has been linked to topics such as mate competition and communication seems rather tenuous. Moreover, to emphasize competition for food as Morse has done, almost at the expense of other exciting areas of behavioural ecology, produces a somewhat imbalanced treatment of the subject. This, then, is a readable and attractively produced book, but one which presents a rather specialized view of behavioural ecology. □

T.R. Birkhead is a Lecturer in Zoology at the University of Sheffield.

A guide to some of the various branches of biology, *Biology in Profile* (editor P.N. Campbell), has just been published by Pergamon. Written especially with aspiring university students in mind, the book contains essays on 20 biological subjects. Prices are: hbk £8, \$19; pbk £3.95, \$9.50.